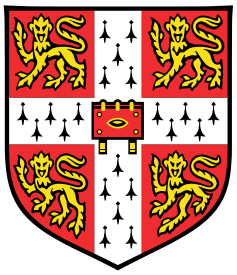
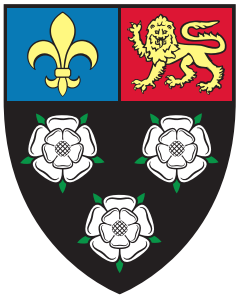


Modelling the self-assembly and structure of carbonaceous nanoparticles



Kimberly Lauren Bowal

Department of Chemical Engineering and Biotechnology
University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

King's College

September 2020

I dedicate this thesis to the one who is over all, through all, and in all.
Lord God Almighty, let my life bring glory to your name.

You keep him in perfect peace
whose mind is stayed on you,
because he trusts in you.
Trust in the Lord forever,
for the Lord God is an everlasting rock.

Isaiah 26:3–4

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification. This thesis is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in Appendix A. The work was performed at the University of Cambridge between October 2016 and September 2020. This thesis contains fewer than 65,000 words including appendices, bibliography, tables, and equations and has fewer than 150 figures.

Some of the work in this thesis has been published:

1. **Bowal, K.**, Martin, J.W., & Kraft, M. (2019). Partitioning of polycyclic aromatic hydrocarbons in heterogeneous clusters. Carbon, 143:247-256

2. **Bowal, K.**, Pascazio, L., Wang, H., Chen, D., & Kraft, M. (2020) Surface properties of heterogeneous polycyclic aromatic hydrocarbon clusters. Proceedings of the Combustion Institute, 38

3. **Bowal, K.**, Grančič, P., Martin, J.W., & Kraft, M. (2019). Sphere Encapsulated Monte Carlo: Obtaining minimum energy configurations of large aromatic systems. The Journal of Physical Chemistry A, 123(33):7303-7313

4. **Bowal, K.**, Martin, J.W., Misquitta, A.J., & Kraft, M. (2019). Ion-induced soot nucleation using a new potential for curved aromatics. Combustion Science and Technology, 191(5-6):747-765

The work presented in this thesis has also contributed, to a lesser degree, to the following publications:

5. Martin, J.W., **Bowal, K.**, Menon, A., Slavchov, R.I., Akroyd, J., Mosbach, S., & Kraft, M. (2019). Polar curved polycyclic aromatic hydrocarbons in soot formation. Proceedings of the Combustion Institute, 37(1):1117-1123

6. Martin, J.W., Botero, M., Slavchov, R.I., **Bowal, K.**, Akroyd, J., Mosbach, S., & Kraft, M. (2018). Flexoelectricity and the formation of carbon nanoparticles in flames. The Journal of Physical Chemistry C, 122(38):22210-22215

7. Pascazio, L., Martin, J.W., **Bowal, K.**, Akroyd, J., & Kraft, M. Exploring the internal structure of soot particles using nanoindentation: A reactive molecular dynamics study (2020). Combustion and Flame, 219:45-56

Kimberly Lauren Bowal
September 2020

Acknowledgements

I'd like to express my thanks to my supervisor Professor Markus Kraft for giving me freedom to explore and putting up with my endless applications to participate in academic activities far and wide. Thank you also to Dr. Jethro Akroyd and Dr. Sebastian Mosbach for keeping the machine that is the CoMo group running smoothly.

Thank you to my collaborators: Alston Misquitta (Queen Mary University of London), Dongping Chen (Beijing Institute of Technology), and Peter Grančič (University of Natural Resources and Life Sciences, Vienna). I am also grateful for the financial support provided by the Cambridge Trust, King's College, and the Sir James Lougheed Award of Distinction.

I was blessed to do my PhD alongside Gustavo Leon, Angiras Menon, and Eric Bringley. Although I still don't quite see the appeal of playing chess online, I've enjoyed going on this PhD journey with you all! Thank you also to Jacob Martin for providing valuable guidance and encouragement, especially during my first year when I was absolutely clueless.

And of course thank you to other CoMo Sapiens and CEB PhDs/postdocs, both past and present (too many to list here). You've helped my time at the department become much more than simply about doing research. In particular, twice-a-day tea times, Thursday morning badminton, and Friday cakes have been a valuable staple of my time here.

My time in Cambridge was made so rich by the relationships I developed with so many unique and wonderful people. I'm grateful especially for friends made through the Christian Graduate Society and King's College.

I'd also like to say thank you to my incredible family: my father, who shows me what it looks like to live with boldness, humility, and integrity; my mother, who I can always count on to provide truthful feedback and motivate me when I am discouraged; and my siblings, who are truly my best friends even when we live so many miles apart.

And, as always, thank you to my Lord and Saviour Jesus Christ for his grace, love, and the ability to learn and discover.

Abstract

The self-assembly and structure of carbonaceous particles are investigated using molecular modelling methods. This provides a deeper understanding of molecular interactions relevant to pollutant formation and growth in combustion processes and other carbon-based applications. The existing soot particle model, a cluster containing planar pericondensed polycyclic aromatic hydrocarbons (PAHs), is extended to include PAHs of varying sizes. The resulting nanostructures show that the classic core-shell morphology reported experimentally for mature soot particles is not energetically feasible if only considering physical interactions between PAHs. It is proposed that young soot particles present the inverse molecular size partitioning. A detailed survey of the surface properties of heterogeneous PAH clusters is conducted, identifying composition-, size- and temperature-dependent behaviours. A novel stochastic global optimisation method, the Sphere Encapsulated Monte Carlo method, is also developed to allow minimum energy structures of large aromatic systems to be determined at considerably less computational expense than existing methods.

The properties of curved PAH molecules are then investigated, and it is hypothesised that their enhanced electronic interactions could play a role in soot particle nucleation. A new intermolecular potential, curPAHIP, is developed to allow the simulation of curved PAHs. Subsequent dynamic clustering studies show that there is a significant increase in particle formation for systems containing curved PAHs and cations, suggesting the importance of these interactions in combustion processes. Further work investigates the structure of clusters containing curved PAHs, and the corresponding influence of cluster size, molecule size and curvature, molecular ratio, and presence of ions.

This work develops computational tools useful for examining large systems of aromatic molecules as well as those containing curved species. Detailed studies on nanoparticle nucleation, structure, and surface properties provide valuable information on self-assembly processes crucial to understanding the production and properties of carbonaceous nanoparticles.

“In God’s wisdom, he determined that the world wouldn’t come to know him through its wisdom.”

1 Corinthians 1:21

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Motivation

The last half century has witnessed remarkable progress in scientific research and discoveries, with breakthroughs in fundamental fields (for example, the mapping of the entire human genome [195] and the detection of gravitational waves [1]) as well as engineering applications (for example, using advanced nanomaterials within electronics [73] and construction materials [248]). The contributions of computational resources to these advancements are difficult to exaggerate. The massive increase in computational power and capabilities has allowed fundamental physical equations to be solved numerically and complex systems to be tackled. Systems at very large and very small length and time scales alike are now able to be explored, predicted, and described.

One system of great interest is carbon-based nanoparticles. Carbonaceous particles too small to see with the naked eye are all around us. Many of these particles serve useful purposes, for example as pigments and material enhancers, in everyday items. However a large number of these particles also contribute to global health and environmental issues. This means that it is desirable to know how best to produce and use these carbonaceous nanoparticles within industrial processes as well as how to best eliminate them or reduce their harmful effects within other environments. In both cases, an understanding of the self-assembly and structure of these important nanoparticles is crucial. This provides information on the role of a particle's composition and conditions on its formation, properties, and behaviour.

This thesis focuses on exploring the self-assembly of carbonaceous nanoparticles at a molecular level using computational tools. The system scope is confined to combustion-produced

particles made up of polycyclic aromatic hydrocarbon molecules (PAHs). Particular attention is spent on examining the role and effects of molecular heterogeneity on particle formation and structure. The first part of this thesis investigates heterogeneous clusters containing planar PAHs, while the second part explores heterogeneous clusters containing curved PAHs. Both parts involve the development of novel computational tools to enable accurate and efficient calculation of these carbonaceous nanoparticles. This provides valuable insight into the properties and behaviours of young soot particles, which are difficult to measure experimentally.

1.2 Novel contributions of this thesis

This thesis provides the following novel contributions:

- The morphologies of heterogeneous PAH clusters containing different molecule sizes are studied to understand molecular arrangements of soot particles. Low energy clusters show a stacked core-shell structure that suggests that young soot particles possess the inverse morphology of mature nanoparticles.
- The detailed surface properties of heterogeneous PAH clusters are explored. The surface composition, roughness, and reactive site densities are related to particle size and temperature.
- A low cost stochastic global optimisation method, Sphere Encapsulated Monte Carlo, is developed for aromatic systems that overcomes existing issues with ring interlocking. This method provides a 10-fold reduction in computational time and is applied to PAH clusters beyond the scope of existing methods.
- An intermolecular potential, curPAHIP, is developed for curved PAHs. This potential allows accurate study of these curved aromatics by incorporating the flexoelectric effect caused by pentagon incorporation.
- The heterogeneous nucleation of planar and curved PAHs with cations is studied, showing that interactions between polar aromatics and ions are able to stabilise small clusters.
- The morphologies of heterogeneous clusters containing curved PAHs, planar PAHs, and cations are investigated. Molecule type and size strongly influence cluster self-assembly and mesophase formation.

1.3 Thesis structure

This thesis is organised into nine chapters, which are structured as follows. After some necessary background information, Chapters 4-8 provide investigations into the self-assembly and structure of carbonaceous nanoparticles, each containing further background and methodology details relevant to the research system under investigation. Chapters 4-6 focus on systems containing planar polycyclic aromatic hydrocarbon molecules while Chapters 7 and 8 examine clusters containing curved polycyclic aromatic hydrocarbon molecules. The final chapter concludes and summarises the findings of the thesis and provides suggestions for further work.

Following this introduction, **Chapter 2** provides background information on the formation and properties of carbon nanoparticles in flames important to understanding the scope of this thesis.

Chapter 3 outlines the theory and methods that are used in this thesis.

Chapter 4 explores the effect of molecular heterogeneity on the internal nanostructure of clusters containing planar polycyclic aromatic hydrocarbon molecules.

These clusters are further characterised with a detailed survey of their surface properties in **Chapter 5**.

A novel Monte Carlo method is presented in **Chapter 6** that allows for the efficient computation of these large and complex systems of planar aromatics.

The development of an intermolecular potential to capture the electronic properties of curved aromatics along with an investigation into the influence of ion-mediated interactions on particle formation are provided in **Chapter 7**.

Further analysis quantifies the self-assembly of homogeneous and heterogeneous clusters containing curved aromatics in **Chapter 8**.

Chapter 9 then provides the conclusions of this thesis as well as suggestions for future work.

Finally, **Appendix A** provides information about the scope of collaboration in this thesis and **Appendices B and C** contain additional details relevant to Chapters 7 and 8, respectively.

Chapter 2

Background

This chapter introduces carbonaceous nanoparticles and provides general context for this thesis with an overview of soot nanoparticles produced within combustion processes. The formation and growth of soot particles as understood through decades of experimental and computational research are reviewed, with a focus on soot particle formation, composition, and structural properties. Further background information and literature reviews are provided in the introduction sections of following chapters to provide a focused context to the research questions addressed there. A brief description of the scope of this thesis is provided to conclude this chapter.

2.1 Carbonaceous nanoparticles

Carbonaceous nanoparticles, defined as carbon-based material at the nanometre (10^{-9} m) scale, are ubiquitous. They are an integral part of everyday life - present in rubber hoses and tyres, plastic cutlery and bin bags, printer ink and electronics, automobile exhaust and baked bread. Carbon nanoparticles are so popular because they possess properties that allow them to serve many useful purposes, such as providing reinforcement, pigmentation, chemical stability, conductivity, and strength. For example, carbon materials are incorporated into optoelectronics, sensors, and imaging probes and used for gas storage, water filtration, drug delivery, and corrosion inhibition. In addition, the production of carbon-based nanoparticles occurs in combustion processes and is desirable in enhancing heat transfer by radiation within industrial furnaces and heat generators. The value of this material is highlighted in its status as the most extensively produced nanomaterial in the world, with demand and production projected to continue growing [103].

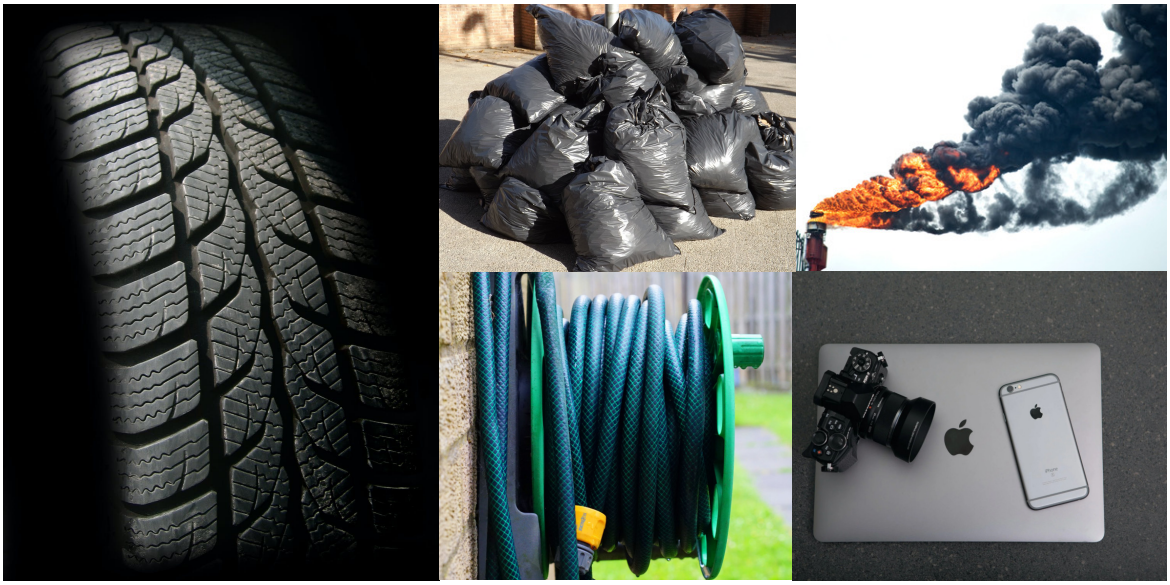


Fig. 2.1 Examples of materials in everyday life that contain carbon-based particles.

However, as with any substance, carbonaceous nanoparticles can also be harmful. Their presence in the air, water, ground, and food [173] has been linked to serious issues in human, animal, plant, and environmental health. Due to their small size and high surface area, carbonaceous nanoparticles present a significant challenge to public health [271, 147, 261]. They contribute to air pollution through reactions with NO_2, SO_2 and O_3 [135, 13], and poor air quality is one of the five leading global health risks (for example, in 2015 it contributed to nearly 8% of all deaths worldwide [72]). These particles are very small (smaller than

dust and less than 1/30th the diameter of a human hair) which allows them to bypass many natural pollutant barriers, such as coughs and sneezes, and enter the lungs and bloodstream, negatively affecting cardiovascular and pulmonary health [204, 164] and causing accelerated aging [201]. Long-term exposure to these particles within polluted air is linked to respiratory infections, chronic obstructive pulmonary disease, stroke, heart attack, and lung cancer [164]. The harmful effects are identified within microorganisms, plants, animals, and humans [2].

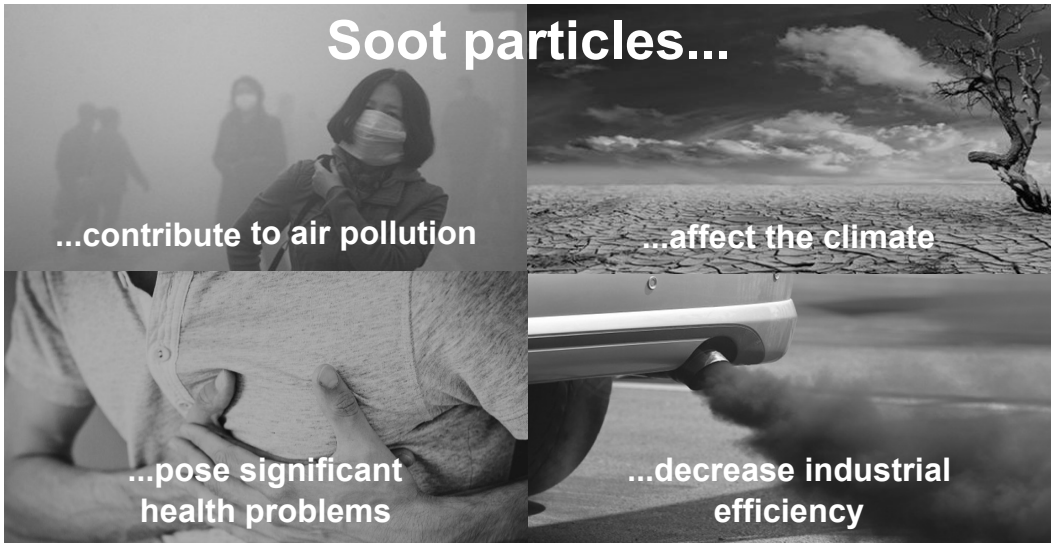


Fig. 2.2 Combustion-produced carbonaceous nanoparticles, known as soot, contribute to harmful effects to the environment, health, and industry.

In addition to their toxic and carcinogenic properties, the molecules within carbonaceous nanoparticles are identified as environmental contaminants [2]. Carbonaceous nanoparticles are considered the second or third major contributor to climate change after CO_2, with an effect of similar magnitude to methane [35, 238]. These dark particles are able to strongly absorb visible light which influences solar radiation with the planet by reducing the albedo effect, especially on large ice surfaces [156, 137, 34, 137, 198]. And, finally, from an industrial perspective, the production of carbonaceous nanoparticles in many combustion processes contributes to decreased efficiency and energy loss [24, 156].

Due to their prevalence and great potential for usefulness or harm, carbonaceous nanoparticles have been studied for decades across many different fields, including materials science, combustion, aerosol science, environmental chemistry, and astrophysics. In this thesis, the focus is on combustion-produced carbonaceous nanoparticles known as soot. This focus shapes the motivating hypotheses and parameters selected, but the fundamental and detailed molecular approach used allows many of the results to be generally applicable to carbonaceous nanoparticle systems beyond the field of combustion.

2.2 Soot particles in flames

The production of carbonaceous particles in combustion systems and their potential for practical use have been known for a long time, probably since humans first interacted with fire. Soot and ash were used to create cave drawings up to 30,000 years ago (see Figure 2.3) [253, 340], and Egyptian documents from 1500 BC describe the use of carbon material to adsorb odours from wounds. The Old Testament itself includes instruction to use ashes for the purification of water in a ceremonial cleansing (Numbers 19:9).

Fig. 2.3 A wooly mammoth cave drawing. This is one of hundreds of drawings left by Paleolithic artists on the walls of the Rouffignac Cave in southern France estimated to be 17,300 years old using carbon dating. Reprinted with permission from Taylor & Francis [340].

However, a detailed understanding and the subsequent intentional manipulation of soot formation in combustion environments has taken much longer to develop. Indeed, this is still an area of intense research, motivated often by the desire to develop predictive soot models to design and control clean combustion systems and efficient flame-synthesised nanoparticle processes. The soot formation process is beautifully complex and the use of advanced experimental and computational tools is required in its elucidation.

In order to provide accessible systems, the majority of research that studies the formation and behaviour of soot particles examines experimental flame systems that reach up to 1600 K and several atmospheres of pressure and that exhibit laminar flow [372]. This controlled flame can typically be divided into several zones, moving from bottom to top: a non-luminous area directly above the burner, the main reaction zone which is often blue/green in colour, a dark space where oxygen is nearly completely consumed, and a final zone extending upwards in which soot particle incandescence produces a yellow/orange glow [36]. This is depicted in Figure 2.4.

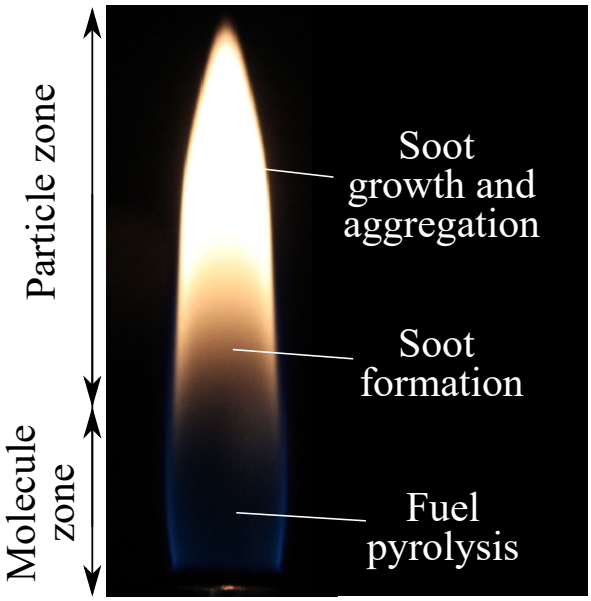


Fig. 2.4 Molecule and particle zones shown on a representative laminar n-heptane/toluene diffusion flame. Flame image modified from [96] and reprinted with permission from Elsevier.

The soot formation process has a number of key stages, highlighted in Figure 2.4: the pyrolysis of fuel to form gas phase molecules, the nucleation of these molecules to form nascent soot particles, and the subsequent growth and dynamics of these particles through chemical and physical processes. This whole process takes place in a very short time span and results in the conversion of hydrocarbon fuel molecules containing few carbon atoms to carbonaceous agglomerates with millions of carbon atoms. For example, it only takes a few milliseconds for gas phase molecules to progress into particles with diameters of 50 nm. Each of the constituent steps is an intricate balance of numerous factors and addressing them fully would (and does) take numerous textbooks and reviews. For the purpose of this thesis, an overview of the key experimental and computational findings and conclusions of the soot formation process will be presented, focusing on the composition and properties of soot particles. If further details are desired, the following excellent books and reviews are worth attention: [32, 125, 226, 76, 352, 146, 182, 270].

2.2.1 Soot precursor molecules

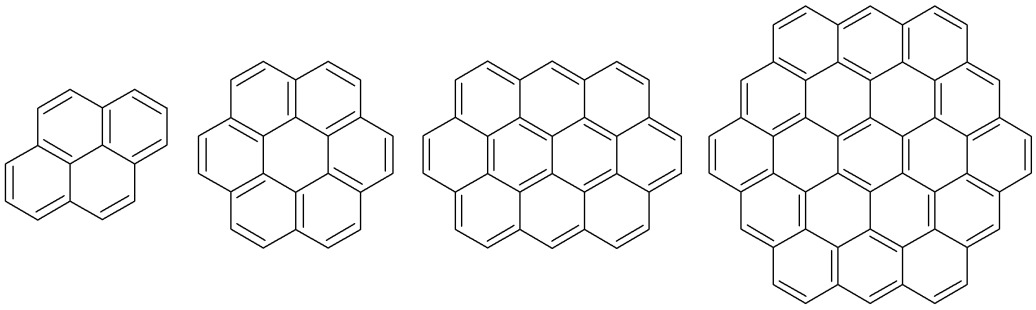
The first stage of soot formation involves the thermal dissociation (*i.e.* pyrolysis) of the long-chain and aromatic hydrocarbon molecules found in fuel to break it into smaller hydrocarbons (containing one to five carbon atoms each) and reactive transient species, including aliphatic

radicals, resonantly-stabilised free radicals, and aromatic radicals [372]. Interestingly, it appears that many petroleum-derived fuels produce the same key pyrolysis products: ethylene (C_2H_4), hydrogen (H_2), methane (CH_4), propene (C_3H_6), 1-butene ($1-C_4H_8$), iso-butene ($i-C_4H_8$), benzene (C_6H_6) and toluene (C_7H_8) [353, 364]. Acetylene (C_2H_2) is another key intermediate that was identified as early as the nineteenth century to participate in flame chemistry [200]. Acetylene is known to play a significant role in the hydrogen-abstraction-carbon-addition (HACA) mechanism [115, 270] in which C_2H_2 groups replace hydrogen atoms, which provides fast step wise growth that drives gas phase chemical reactions and grows the pyrolysis products into larger stable cyclic molecules.

Extractive and *in situ* experimental methods such as electron microscopy, laser-induced incandescence, and laser-induced fluorescence are used to identify the gas phase species present in the flame reaction zone and investigate the reaction kinetics leading to soot particle precursor molecules [375]. There are a large number of species present within the gas phase of a flame. Experimental and computational work has focused on identifying the dominant species as well as those most influential to the soot formation process. Laser analytical techniques have found that molecules with an average C/H ratio of 1–5, mass of 450–650 Da, and size of 1.0 ± 0.2 nm are present in significant quantities [7, 148]. 35 years ago, Stein and Fahr [307] analysed the stability of hydrocarbon molecules at high temperatures and identified a class of molecules called stabilomers, comprised mostly of pericondensed polycyclic aromatic hydrocarbon molecules (PAHs) containing an even number of carbon atoms, that present the most stable bonding configurations of hydrocarbon molecules within combustion environments. The dominance of these stable PAH molecules in the gas phase of combustion environments was supported by experimental and computational studies [281, 39, 231, 229]. For example, gas chromatography - mass spectrometry analysis shows that the gaseous phase before soot formation is dominated by PAHs containing four or fewer rings [235]. In fact, PAHs are found in all sooting and hydrocarbon flames [149]. Recent work using laser desorption ionisation, coupled with high resolution mass spectrometry and imaging analysis identified that the majority of species observed consist of fused six-membered rings [229]. PAHs have thus been widely accepted as important intermediates in soot formation due to their stable bonding configuration at flame temperatures [100, 149, 307, 87] and are even proposed to contribute a significant fraction of the carbon matter present in the interstellar medium [33].

Symmetric pericondensed PAHs, such as those shown in Figure 2.5, are used as the fundamental building block within chemical mechanisms and soot models, however, as discussed further in later sections, they cannot completely explain the behaviour and properties of soot observed. This has caused an increasing focus in recent years on other molecule types

within the gas phase that may play a significant role in soot formation and growth. This has focused on polycyclic molecules that are curved due to the presence of non-hexagonal rings, PAHs bonded with covalent cross-links, and resonantly-stabilised radicals. Laser induced fluorescence spectroscopy shows signals above 550 nm that cannot be explained only by peri-condensed PAHs [23], suggesting that PAHs with five-membered rings or covalently linked small PAH dimers are very likely to be present. Computational studies using density functional theory, transition state theory, and kinetic modeling (0D reactor simulations) show that the growth of five-membered rings in large PAHs is very likely [208]. Aberration-corrected TEM and electron energy loss spectroscopy have suggested the presence of pentagonal carbon rings [14] and recently high resolution atomic force microscopy has provided direct images of potential soot precursor molecules, showing that molecules containing five-membered rings and covalent cross-link bonds are also present in the flame gas phase [291]. Kinetic Monte Carlo simulations of gas phase molecule growth show a great diversity in aromatics, with different degrees of oxygenation and aromaticity, including five-membered rings, furan, hydroxyl, ketone groups [283]. Stochastic simulations and aerosol mass spectrometry also find an abundance of PAHs containing five-membered rings and substituted aliphatic chains that could contribute to the formation of young soot particles [186].



(a) pyrene ($C_{16}H_{10}$) (b) coronene ($C_{24}H_{12}$) (c) ovalene ($C_{32}H_{24}$) (d) circumcoronene ($C_{54}H_{18}$)

Fig. 2.5 Examples of symmetric pericondensed PAH molecules.

The growth of PAHs and other species in flames is controlled by kinetics [352] and significant work has been done to develop and apply chemical kinetic mechanisms of gas phase reactions within flame environments. These kinetic models have been increasing in size and number over time and today contain up to 100+ species and 1000+ reactions, often built in a ground-up approach to understanding soot particle formation [212, 115, 67, 264]. This is an area of research that is still progressing and the detailed chemical mechanisms and the identity of all relevant chemical species are still being developed and better understood.

2.2.2 Gas-to-solid particle formation

Mechanistic models describing soot formation often split the process into two parts: the gas phase chemistry and the soot particle dynamics. The intermediate step in which gas phase species transition into solid phase soot particles is a vital part of soot formation but is oversimplified or non-physical in existing models. This is because the gas-to-solid transition mechanism is poorly understood, primarily due to the fact that it occurs in a very small time frame and length scale which make experimental observations nearly impossible. Many techniques do not capture the same sizes (for example, gas chromatography is typically able to detect species < 300 amu while HRTEM can measure species > 1.5 nm [270]) and conventional aerosol instruments, such as condensation particle counters, face difficulties in characterising particles below 2 nm [30]. As the least well understood step within the soot formation processes, soot particle formation has accordingly been a focus of intense research for many years, with the most recent investigations centring around the role of chemical versus physical interactions.

Chemical soot inception

Inception typically refers to the hypothesis of soot formation as a primarily chemical process involving the formation of covalent bonds between molecules to form solid particles. Early hypotheses proposed that soot is formed through the chemical polymerisation of acetylene molecules [262] although these were discredited because the required chemical reaction rates are too slow to account for the observed rate of soot production. Another suggested mechanism involved the formation of icospirals, fullerene-like molecules, that act as a seed for soot particle growth. This was supported by electron micrographs of soot particles where closed and spiral fringes indicated a fullerene-like seed [332] but these have since been attributed to overexposure to electron irradiation during experimental analysis [347, 333]. In addition, the fraction of hydrogen measured using molecular beam mass spectrometry supports the view that PAH growth does not extend past 1000 Da in most sooting flames, which would be required for icospiral formation. Further work suggested that polyyne molecules, observed in significant concentrations in flames [158], may be a key precursor for nanoparticle formation, but this is unlikely since their rigidity makes them unlikely to cyclise and they are prone to degrade in flame environments [207]. Another polyyne mechanism in which fullerene cage formation occurs has been proposed in recent years with the use of reactive force field simulations [136, 368]. However this mechanism also presents significant issues, including a requirement for very high pressures/concentrations, a slow reaction rate

that is limited by hydrogen abstraction, a high C/H ratio, and the presence of long polyynes not seen experimentally [214].

Further work has identified the potential influence of cross-linking covalent reactions between PAHs [115, 77, 339], which produces particles that have hardness and strength values in agreement with experimental results [251] and may reflect the nanostructure observed in graphitised mature particles. Importantly, these cross-linked aromatic species possess bonding energies that are high enough to explain their presence at flame temperatures [213]. The presence of aliphatic side chains on condensed-ring structures can strengthen interactions between colliding molecules so they stay close enough together for reactions to occur [102]. The stabilising effect of side chains is not universal, however, with short flexible chains shown to increase dimer stability but longer chains causing steric hindrance that reduces stability. It appears that the presence of side chains is also not enough to compensate for destabilisation of a four-ringed PAH dimer at 1000 K, indicating that other stabilising factors are required. Although some cross-linking interactions may play a significant role in soot particle formation and/or growth, there are significant issues with existing mechanisms that need to be further explored, including the fact that model particles composed of cross-linked PAHs possess higher quantities of carbon than expected when compared to the C/H ratio identified in soot particles, the requirement of a large number of radicals within this mechanism, the rate of formation of cross-links is too slow to completely explain the observed production of soot particles in flames, and the molecules are prone to radical-induced fragmentation [115, 355].

Physical soot nucleation

The second major approach to understanding the gas-to-solid transformation step in soot particle formation evaluates the role of physical interactions and often uses the term nucleation to refer to soot particle creation. Following the understanding that soot precursor PAHs interact predominantly through dispersion interactions, the general assumption is that the PAHs grow within the gas phase to around 1–2 nm, at which size they are able to come together through a collision interaction that allows attractive forces to form an initial solid soot particle. This physical mechanism follows from the large attractive potential between adjacent molecular sheets (5 $kJmol^{-1}$ / C atom) which allows π-π interactions to contribute to condensed phase structures throughout nature such as protein and DNA structure.

Mass spectra of fragments from photoionisation of very young soot particles show periodicity, suggesting that the mass of very young particles increases discretely when growing and therefore interactions between independent molecules to form a dimer, trimer, etc. are

crucial to soot nucleation and growth [140]. Extractive and *in situ* optical measurements also suggest that PAH molecules grow through gas phase reactions until they are sufficiently large that they agglomerate into ordered stacks [7, 39]. This mechanism of dimerising through physical interactions also explains the rapid formation of nuclei in low to intermediate temperature flames with a collision factor of 10^{-6} to 1 and is supported by the experimental evidence for PAH stacking in soot particles [20]. Computational calculations of these physical interactions have shown that the binding energies of dimers containing stabilomers increases with molecular size and dimers with reduced masses as low as 83 Da are able to exist at 1500 K [149]. Several studies have postulated that the conversion of translational energy of colliding PAHs to rotational energy with the development of internal rotors may have a stabilising effect and permit particle formation through physical dimerisation and subsequent chemical growth [230, 290].

The hypothesis that PAHs assemble into particles through van der Waals attractive forces has received significant attention and is the nucleation method assumed in most soot growth models. In particular, following the initial use of PAH dimerisation as the nucleation step within a soot formation model by Frenklach and Wang [120], the four-ringed PAH pyrene (shown in Figure 2.5) has been adopted as the critical molecule size in order to provide agreement with observed nucleation rates and particle size distributions. Pyrene dimerisation is an attractive pathway to particle nucleation and has been widely implemented as the soot formation mechanism within computational models of combustion ([99, 265, 17] as a few examples). However, many experimental and computational studies have called this into question, investigating whether the relatively weak van der Waals interactions between small PAHs such as pyrene are sufficient to explain the formation of soot particles at the high temperatures within flames. Kinetic and energetic information from experimental measurements and theoretical calculations show that physical dimerisation of pyrene molecules via van der Waals forces is not a likely pathway to soot nucleation since the equilibrium of the dimerisation strongly favours dimer dissociation at the high temperatures of combustion environments [279]. Using well-tempered metadynamic simulations, dimer stability was found to be dependent on monomer mass, with heterogeneous dimers reflecting the stability of the size of the smallest monomer [211]. The infeasibility of PAH condensation at high temperatures was also concluded with independent all-atom dynamic and reactive simulations of PAH clustering [327, 71, 214]. The reversibility of these dimer interactions is also found to be significant when investigated with a kinetic mechanism, such that irreversible chemical bond formation is required to reach the much higher soot nuclei concentration observed with laser induced incandescence measurements [186]. Mass spectrometry and photoionisation studies combined with stochastic chemical kinetics simulations suggest that the dominant

species observed at mass 202 u using aerosol mass spectrometry are not pyrene but and instead are other non-stabilomer (*i.e.* not even carbon numbered, unsubstituted) $C_{16}H_{10}$ isomers [176]. It is now generally accepted that the dispersion forces between two small PAHs such as pyrene are not significant enough to solely account for observed soot formation. Although it may be thermodynamically possible for larger PAHs (>20 rings) to dimerise through physical interactions alone at combustion-relevant temperatures, these molecules are not present in significant concentrations in the soot-forming region of flames. In addition, the mechanical strength of particles containing PAHs only bound by physical interactions does not match the measured hardness of soot particles [251]. As a result the exploration for a replacement precursor molecule(s) and mechanism(s) continues.

This is an area of ongoing investigation and requires further work to elucidate the pivotal soot particle precursors and formation mechanisms. In particular, the balance between thermodynamics and kinetics is important. It is increasingly clear that no single mechanism (for example, the chemical or physical interactions discussed above) can fully account for the formation of soot particles. There is a growing body of experimental and computational research suggesting that radicals, ions, and/or molecules containing five-membered rings may play a role in soot formation [176]. In particular, PAHs containing embedded five-membered rings possess curvature and polarity that may allow them to interact significantly with ions. This will be introduced and explored further in this thesis.

2.2.3 Soot particles

There are also many unanswered questions concerning the detailed composition and structure of freshly formed young soot particles and how these properties change as the particles mature. Soot particles can be classified into two stages: young particles with a diameter of a few nanometres and mature particles with a mobility diameter greater than 10 nm.

Nascent soot particles

Young soot particles, also known as nascent particles, precursor particles, or elementary particles, are the smallest nanoparticles measured in flames. For the same reasons that soot particle formation is not yet fully understood, namely the experimental challenges in detecting, capturing, and analysing small particles (< 3 nm) at flame conditions as well as the complexity of accurate numerical models, young soot particles are not well characterised.

Although the existence of small nascent soot particles of 1.5–2 nm was observed in flames in 1973 by Wersborg et al. [358] using molecular beam sampling, experimental work continuing the investigation of nascent soot particles did not progress significantly until nearly two decades later when two groups independently detected particles only a few nanometers in size using *in situ* [75] and *ex situ* [84] techniques. Since then further advances have provided a plethora of data using many techniques including differential mobility analysis [297], small-angle X-ray scattering measurements [83], and flame-sampling photoionization mass spectrometry experiments [134].

It appears that the smallest soot particles contain PAH molecules and are approximately 2000 amu with an effective diameter of 1.5 nm [270, 319]. These young soot particles are highly reactive [52], do not absorb visible light [228], have C/H ratios of 1.4-2.5 [277, 31, 75], densities between 1.12–1.50 g/cm^3 [323, 51, 85], and possess more aromatic character than aliphatic [50, 247]. Helium ion microscopy and atomic force microscopy show nascent soot as liquid-like transparent particles without a well-defined shape [287, 54, 3]. Transmission electron microscopy image analysis shows that nascent soot particles typically possess an amorphous-like internal structure with short curled fringes [16], although there appears to be a tendency for molecules around the edge of the particle to be slightly larger and more stacked than those in the core [41, 81]. It is important to note that these experimentally-observed soot particles are likely kinetic products rather than thermodynamic ones.

Mature soot particles

As soot particles spend more time in the flame, their chemical and physical properties change and they transition into mature particles. In addition to growing in particle size, the maturation process causes constituent molecules to increase in length and decrease in intermolecular spacing and tortuosity, and thus form more ordered stacked arrangements [16, 350, 41], indicating carbonisation. These mature particles are solid, spheroidal, absorb in the visible and UV range, possess densities of 1.4-2 g/cm^3, and show C/H ratios between 3–20 [81, 28]. X-ray analysis indicates the presence of a regular graphitic structure throughout mature soot particle volume and electron microscopy analysis shows that mature soot particles are typically composed of molecule stacks, with a core-shell nanostructure in which the shell contains larger molecules with lower tortuosity and a higher stacked percentage (*i.e.* a higher degree of orientation) than the core [18, 16, 40].

2.2.4 Soot aggregates

Mature soot particles grow via coalescence and surface reactions with gas phase species, including pyrolysed fuel species and PAHs [270]. Upon reaching about 10-20 nm in diameter, they also undergo coagulation and agglomeration by particle-particle collisions [188]. The aggregation of two or more mature particles form multicore particles, followed by additional growth and oxidation cycles [79, 270]. Through coagulation, nearly spherical primary particles form chain-like structures with a fractal dimension of around 1.8 [86]. Other competing events, such as further dehydrogenation, surface and internal oxidation [306, 366], sintering, and breakages can also cause the fractal-like particle agglomerates to shrink. These have been studied in detail using population balance modelling tools [57, 205] as well as particulate characterisation techniques [357, 337]. All of these processes ultimately produce large soot particle agglomerates made up of mature soot particles with a total radius of gyration on the order of 100–400 nm [10]. These agglomerates then leave their combustion system and are often released into the surrounding environment such as the air, soil, and water.

The dynamics of soot particle populations are often modelled using the Smoluchowski coagulation equation within a discrete sectional method (where particle properties are averaged within each section of particle ensemble) [260] or method of moments (which uses moments of the particle size distribution function) [117] to provide insight into the evolution of the particle ensemble. Components of the whole soot formation process including fuel pyrolysis, PAH formation, nucleation, surface growth reactions, oxidation, and particle processes, have been implemented in both method of moments [114] and sectional models [356, 18]. Detailed evaluation of soot surface growth can be performed with dynamic Monte Carlo simulations [113, 199, 359]. Computational fluid dynamics simulations also allow larger system properties such as pressure and fluid flow to be incorporated into soot particle modelling [9, 348].

2.3 Measurement weaknesses

Experimental methods have been crucial in shedding light on the chemical and physical behaviours of hydrocarbon molecules and particles within combustion environments, but they still have limitations that are worth noting. Most observations about the internal structure of a soot particle (particularly the quantification of constituent molecule curvature) are based on TEM images, which are determined from a two-dimensional projection of the condensed

three-dimensional nanoparticle in which only layers parallel to the incident electron beam can be imaged. It is not well quantified and understood to what extent the intrusive sampling methods such as thermophoresis influence the particles being sampled. For example, it is believed that the highly ordered carbon onion structures characterised by concentric graphitised shells that were imaged in early studies [166, 333] were caused by the irradiation of samples in electron microscopy which caused carbonisation [160, 347]. This is still likely a factor in similar studies conducted since then. In addition, flame disturbance caused by intrusive instruments and thermocouple temperature reduction due to soot deposition are poorly defined factors. Another crucial weakness is that experiments are predominantly focused on mature soot particles. Optical methods such as laser induced incandescence are only sensitive to mature soot particles and cannot measure the temperature or soot volume fraction in regions dominated by liquid-like nascent soot particles [185]. Nascent soot particles exhibit sizes of a few nanometres and might possess electric charges, complicating their characterisation with differential mobility analysis. X-ray diffraction, commonly used to obtain crystal structures including interlayer and intralayer distances and crystallite size [202], cannot provide an unequivocal identification of structure (for example, it cannot show whether pentagonal or hexagonal rings are present). Experimental measurements need to be more sensitive, accurate, noninvasive and able to well sample the physical and chemical characteristics of soot during development and evolution [226] - these issues can be partially addressed using computational studies.

2.4 Computational studies

One valuable way to further investigate and understand both the formation and characteristics of nascent and mature soot particles is to develop and use soot particle models within a computational framework. A number of approaches have been used to do this, varying in their representation of soot particles as well as their computational methodology. Coarse-grained models represent soot particles as small systems of rigid elliptical components, each representing a constituent molecule, similar to preceding work studying liquid crystal structures [151, 163, 153]. Other methods directly populate the particle species using the fringe lengths identified in experimental TEM images [107] or using the composition and mass of molecules identified using kinetic Monte Carlo simulations [338]. A dominant soot particle model uses clusters containing pericondensed PAHs (such as those identified as stabilomers within the range of fringe lengths observed in TEM images) at atomic resolution [149, 323, 290, 165, 267]. Stochastic energy minimisation methods, reactive and

non-reactive molecular dynamics, and kinetic Monte Carlo are the most common tools used to examine these model systems. These studies have provided valuable insight into the energetic and spatial properties of proposed soot particle systems, for example by elucidating information about particle melting points [64, 61, 37], surface interactions [132, 60], and molecular arrangements [267, 323, 153].

Existing soot models have largely ignored the effects of heterogeneity in soot particles and paid little attention to the internal structure of these systems, especially for young particles. This is a significant weakness since the nanostructure of soot particles plays a crucial role in determining how soot primary particles form and grow and how larger aggregates form and interact. Therefore there remain many unanswered questions, such as the role of different molecule sizes and shapes, and other components such as cations, within young and mature soot particles. The work presented in this thesis seeks to address many of these points through carefully expanding a soot particle model to include other experimentally observed components and characterising the particle nanostructure using computational techniques. Further details on the status and limitation of soot particle studies will be provided within the introductions of further chapters to provide the context of the research questions addressed.

2.5 Scope of thesis

There is a plethora of experimental data on the physical and chemical properties of soot particles, however the computational models and fundamental bottom-up understanding of these systems remain poor, especially for nascent soot particles. Detailed molecular models investigating the nucleation and properties of soot particles remain simple, considering only homogeneous planar PAHs. This thesis seeks to address these limitations through the detailed analysis and investigation of heterogeneous systems, first considering the influence of varying PAH size on the internal and surface properties of soot particles. Then the potential role and influence of curved PAH molecules containing five-membered rings is explored within the context of homogeneous and heterogeneous soot nucleation. Finally, the self-assembly and structure of nanoparticles containing curved PAHs is investigated and quantified.

The scope of this thesis is to investigate the self-assembly of carbonaceous nanoparticles relevant to combustion processes. This thesis addresses the formation of soot particles from gas phase PAHs and the internal and surface structural properties of these nanoparticles at an atomic level. This is done through the extension of existing molecular models of carbonaceous particles to include molecule size heterogeneity, as well as the influence of molecule curvature and the presence of charged species. A number of numerical techniques

are used to produce relevant soot particle models and evaluate their structure and morphology. These advances reflect experimental developments in the known composition of soot particles and require the parametrisation of a new intermolecular potential to capture the behaviour of relevant molecules.

Chapter 3

Computational Methods

This thesis investigates the behaviour and properties of carbonaceous soot nanoparticles using molecular dynamics and Monte Carlo simulations. This chapter provides an overview of the theory behind these molecular modelling tools, highlighting their respective strengths and weaknesses. Some methodology and theory details which are specific to individual chapters are omitted or only briefly included here and will be explained later in their respective contexts.

In order to properly model and examine molecular systems, two components are required: a model that accurately describes the interactions between atoms and a rigorous method to evaluate these interactions under different conditions. Interactions are typically captured by interatomic energy potential descriptions while the stable structures and dynamic behaviours of systems are often determined by molecular modelling techniques such as the Monte Carlo minimisation and molecular dynamics. These crucial components will be introduced below.

3.1 Intermolecular potential

Intramolecular forces are those that hold atoms together to form molecules, characterised by different chemical bonds. Intermolecular interactions are those between non-bonded molecules. Both types of molecular force are involved in the self-assembly of carbonaceous soot nanoparticles and it has been proposed that the interactions between non-reacting soot precursor molecules play a key role in the nucleation and early formation of soot particles. Intermolecular forces are weaker than chemical bonds, but become significant for many condensed materials. For example, although the typical non-bonded interactions between small hydrocarbons are weak ($<$ 2 kJ/mol) the binding energy between two 100-carbon graphite sheets (500 kJ/mol) is stronger than a typical single carbon-carbon bond (400 kJ/mol) [149]. The exploration and characterisation of intermolecular interactions has been well studied and is described in detail elsewhere [308]. A brief background and discussion of aspects relevant to the study of soot formation are included here.

Theoretical and computational models to investigate molecular systems require an accurate description of intermolecular interactions. This is done through the use of an intermolecular potential, also known as a force field. A force field is used to calculate the potential energy of a molecular system and is defined by its functional form and its associated parameters. In contrast to *ab initio* calculations, which use quantum mechanics, simulations using force fields are relatively fast and able to handle large systems, such as proteins or young soot particles. Force fields typically consist of four energy contributions: electrostatic, repulsion, induction and dispersion. The electrostatic interaction depends on the charge distribution and can be accurately described, for example using distributed multipoles. The induction term describes the response of the molecular charge distribution to external fields and requires polarisabilities. The dispersion term includes van der Waals attraction. The basic form of an additive force field, shown in Equation 3.1, consists of the sum of the bonded and non-bonded contributions to the system energy.

$$U_{\text{total}} = U_{\text{bond}} + U_{\text{angle}} + U_{\text{dihedrals}} + U_{\text{electrostatic}} + U_{\text{van der Waals}} \tag{3.1}$$

Energy contributions from interatomic bond stretching (U_{bond}), angle bending (U_{angle}), and proper and improper dihedral torsions ($U_{\text{dihedrals}}$) make up the intramolecular components, while the electrostatic ($U_{\text{electrostatic}}$) and van der Waals ($U_{\text{van der Waals}}$) energies comprise the intermolecular components. Many potential forms have been used to describe these interactions and have been parametrised for systems of interest. It has been shown that the potential used has a significant impact on the simulation results [157, 252] and therefore force field selection is important.

One of the simplest and most commonly used force field forms is the Lennard-Jones potential [177], given in Equation 3.2.

$$U_{ij} = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{R_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}}\right)^{6}\right] \tag{3.2}$$

where ε_{ij} is the potential well depth, σ_{ij} is the distance at which the interatomic potential is zero and R_{ij} is the separation distance between atom i in molecule I and atom j in molecule J. The R^{-6} term describes long-range attraction, while the R^{-12} term describes electron repulsion at short ranges. As seen in Figure 3.1, the repulsive energy is a steep barrier when the distance between the atoms is small. Although there is no theoretical justification for the repulsive term and it is known that the LJ potential generates overly repulsive interactions at a short range, this potential remains popular due to its computationally simple form. It has been used often in molecular modelling of PAH interactions [149, 71, 267, 266, 327, 123]. More physically motivated intermolecular potentials have been developed, including the Buckingham potential [47], Morse potential, and Born-Mayer potential [38]. The isotropic PAH potential (isoPAHAP) developed by Totton et al. [324] is fitted to quantum calculations of PAH interactions. Figure 3.1 shows the potential energy curve of carbon-carbon interactions using the isoPAHAP and LJ force fields, and it is apparent that the LJ potential has a significantly larger well depth. This results in an overestimation of the binding capacity of PAHs and has been shown to greatly impact the modelling of PAH dimers and clusters [327].

The isoPAHAP potential has the following form:

$$U_{ij} = G\exp\left[-\alpha_{ij}\left(R_{ij} - \rho_{ij}\right)\right] - f_6\left(R_{ij}\right)\frac{C_{6,ab}}{R_{ij}^6} + \frac{q_i q_j}{R_{ij}} \tag{3.3}$$

where the first term describes the short-range interactions using the Born-Mayer form. The other two terms describe the long-range interactions as the dispersion multiplied by the Tang-Toennies damping function [318], $f_6\left(R_{ij}\right)$, and a point charge electrostatic term. U denotes the interaction energy between atom i and atom j, G sets the energy unit of this term and is taken as 0.001 Hartree, α is the hardness parameter in the Born–Mayer term, R_{ab} is the atom-atom separation where a and b denote atomic sites within a molecule, ρ is a shape parameter, $C_{6,ab}$ is the isotropic dispersion coefficient, and q_i and q_j denote the partial charge on atoms i and j, respectively. The isoPAHAP parameters were fitted using symmetry-adapted perturbation theory density functional theory calculations, and the results are shown in Table 3.1. Further details regarding the development and parametrisation of this force field are described by Totton et al. [324, 327].

Table 3.1 Parameters of isoPAHAP in atomic units [327].

atom pair	ρ	α	C_6
C C	6.0434	1.8783	30.282
C H	4.9562	1.7560	12.605
H H	4.1195	1.4043	5.2179

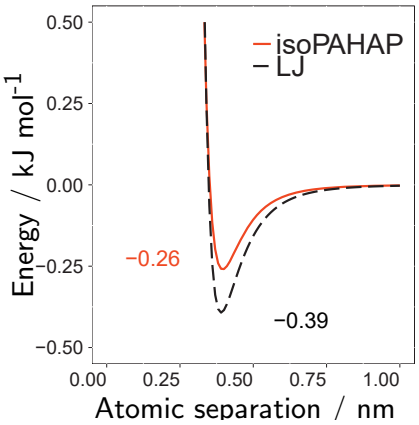


Fig. 3.1 Comparison between the isoPAHAP and LJ potentials for carbon-carbon interactions. The LJ potential parameters are taken from [334]. Figure from [59].

3.2 Molecular modelling

Molecular modelling is a well-established field and there are many comprehensive resources available that describe it in detail [196, 112]. A brief overview of important concepts related to the stochastic energy minimisation (Monte Carlo) methods and the deterministic molecular dynamics simulations involved in this thesis is provided.

In order to understand the energy of a chemical system, the concept of a potential energy surface (PES) is employed. The PES is a hypersurface defined by the potential energy of a system over all possible atomic arrangements. Using the analogy of a physical landscape, the PES provides a complex landscape in which the valleys correspond to states of low potential energy and the peaks correspond to states of high potential energy. Both local and global minima are points of interest on the PES since they correspond to optimised molecular structures. Saddle points, defined as points with no slope, downward curvature in one direction, and upward curvature for all other directions, provide insight into transition states since they lie on the lowest energy barriers connecting minima. Knowledge of all of the minima and saddle points on a PES therefore provides a complete description of all possible chemical structures and connecting isomerisation pathways for a specific system.

3.2.1 Energy minimisation

Mapping out the complete PES of a chemical system is an ideal that is, unfortunately, unfeasible for most systems of interest because of the immense amount of computational effort required to sample the configurational space to this extent. Many methods have been developed to identify the points of interest (minima) as accurately and quickly as possible. No minimisation method exists that can always find the global minimum in a reasonable amount of time, but there are a number of iterative methods that are able find the nearest local minimum from an initial position. The simplest of these methods are function evaluations, where each step is made based on the results of the previous evaluations. The simplex method is one such example. A second type of minimisation method uses derivative information. These methods are most commonly used in molecular dynamics because the partial derivative of the potential energy is already known. Examples include the steepest descent method (which always moves in the direction of the most negative gradient) [56], conjugate gradient method (which also moves down negative gradients, but uses gradient information from previous steps) [155], and the Broyden-Fletcher-Goldfarb-Shanno method

(which is a quasi-Newtonian method that uses an approximated Hessian matrix of second derivatives) [48].

3.2.2 Basin-hopping Monte Carlo method

In order to rigorously locate minima on a complex PES, especially when aiming to determine the global minima, further optimisation and selection methods, within which these minimisation methods can be used, are required. The basin-hopping method is an unbiased Monte Carlo (MC) minimisation method that simplifies the search of a complex PES by transforming it to a series of low energy plateaus, called basins [344]. This is done by taking the energy of each configuration as the energy of a local minimum. As seen in as a one-dimensional representation in Figure 3.2, this essentially maps the PES into discrete energy levels that serve as a series of plateaus around each local minimum. This hyperspace deformation of the PES removes the energy barriers associated with transition states but keeps all minima unchanged.

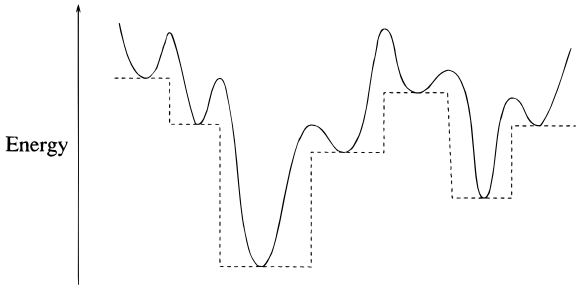


Fig. 3.2 One-dimensional schematic illustrating the energy transformation used in a basin-hopping simulation. The solid line shows the energy of the original PES and the dashed line shows the transformed PES. Reprinted with permission from [344].

After the calculation of each minimised system, all of the molecules are randomly rearranged and another minimisation is conducted from the resulting geometry. This is done in order to allow for the searching method to move between basins on the transformed PES. Implementing a large enough stochastic rearrangement prevents the system from being trapped in a local minima and promotes scanning of the entire PES. If the energy of the new minimum is lower than the starting energy, the configuration of the new minimum is accepted as the next initial configuration. If the energy is greater, a Monte Carlo accept-reject scheme is used to decide the outcome. The basin-hopping method has been shown to be a very effective global optimisation technique since it allows for the determination of the global minimum through

exploration of a small fraction of the existing local minima, and has successfully determined the minimum configurations of many LJ homogeneous and heterogeneous clusters [19]. Further information about the basin-hopping method and its use within PAH systems is provided in Chapter 6, which introduces a new stochastic method to study large PAH clusters.

3.2.3 Molecular dynamics

Molecular dynamics (MD) is a computer simulation method used to study the physical movements of atoms and molecules over time. Pioneering work in the 1950s studied simple systems of interacting hard spheres. The use of molecular dynamics to study materials began with the work of Rahman [263], who introduced the methodology, language, and technology for the simulation of condensed matter that is still used now. The subsequent development and increased capability of computing machines has allowed this field to grow and it is now used to model many systems at the nanoscale, with applications in many fields, including material science, biology, and combustion.

The fundamental principle of MD is simple: the motion of a system of atoms is solved iteratively through time using classical mechanics. This allows equilibrium and transport properties to be determined for a many body system. At each time step the potential energy is calculated from the force field. Newton's equations of motion allow the change in the potential energy of each atom in the system to be expressed in terms of the change in atomic position, as shown in Equation 3.4.

$$-\nabla\vec{U} = \vec{F} = m \times \vec{a} = m \times \frac{\partial^2 \vec{r}}{\partial t^2} \tag{3.4}$$

where U is the potential energy of the atom, F is the force exerted on the atom, m is the mass of the atom, a is the acceleration of the atom, and r is the atomic position. In this way, the movements and interactions of up to tens of thousands of atoms are simulated as a function of time.

The purpose of MD simulations is usually to determine the time evolution (also known as a trajectory) of the molecular system after it is subjected to a perturbation caused by thermal motion and/or external forces. Within MD simulations, this perturbation is often implemented as a velocity distribution among the atoms. A continuous trajectory containing the atomic positions over the duration of the simulation is generated by integrating the equations of motion of the system. This is calculated by an integration algorithm that considers the simulation as a finite number of discrete time steps. Several integration algorithms exist and

one commonly used method is the velocity Verlet algorithm since it is easy to implement and provides accurate integration with temperature coupling. Further information on the three step procedure used to calculate the system velocities and positions are described by Verlet [336].

A thermostat is used to modulate the system temperature in a MD simulation. There are a variety of thermostat methods available, each working by adding and removing energy to the system. Despite being called thermostats, these methods do not simply seek to maintain a constant temperature through a fixed kinetic energy, instead they ensure that the desired average system temperature is maintained. Some popular techniques include the Anderson, Berendsen, and Nosé-Hoover thermostats. The Anderson thermostat uses uses stochastic collisions acting on random particles to represent coupling the system to a heat bath [12]. However, since the post-collision particle momenta are randomly selected from a Boltzmann distribution, the calculated dynamics are not physical, and therefore this thermostat is not desirable if measuring dynamical properties. The Berendsen thermostat corrects deviations of the actual temperature from the desired temperature by rescaling the velocities at each time step [26]. This allows the system temperature to relax quickly but results in a suppression of kinetic energy fluctuations, which means the Berendsen thermostat is useful for equilibration but cannot be mapped to a specific ensemble. Constant temperature MD simulations can be achieved through the use of an extended Langrangian; that is, a Lagrangian that contains added artificial coordinates, potential energy, and momentum related to the real values by scaling. A thermostat using this was developed by Nosé [242] but the most commonly used formulation is from Hoover [159], and so it is called the Nosé-Hoover thermostat. The Nosé-Hoover thermostat successfully simulates a system with energy fluctuations at a constant temperature. Nosé-Hoover chains are required in many cases since a single thermostat is not able to probe a canonical ensemble if there is more than one constant of motion (indicated by a simulation with no external field, an unfixed centre of mass or more than one conservation law).

3.2.4 Replica exchange molecular dynamics

Complex systems are difficult to simulate since their energy landscapes contain many minima. It is often difficult to cross energy barriers between minima at low temperatures within a reasonable simulation time and so the results can be highly dependent on the choice of initial conditions. Replica exchange molecular dynamics (REMD) is an advanced form of molecular dynamics used to improve sampling of the phase space in an MD simulation and avoid becoming trapped in a local minimum on the PES. Also known as parallel tempering,

this method was developed in 1986 for simulations of 'systems with quenched random interactions' [314] and implemented first within MC methods [161] and then adapted for use in MD [313]. In an REMD simulation M copies of the original system, called replicas, are run in parallel. Each replica, usually in the canonical ensemble (defined by maintaining a constant number of atoms, system volume, and temperature), is simulated at a different temperature. High temperature systems, which possess higher entropies, are generally able to sample the phase space more widely, while low temperature systems may be trapped in local energy minima for the duration of a typical simulation. REMD allows the replicas to exchange atomic coordinates or temperatures at certain time intervals. This allows low temperature systems to access more of the phase space. The exchange of atomic coordinates or temperatures is dependent on the probability that the coordinates or temperature states exchanged would have been observed in the two ensembles. This probability is dependent on replica temperatures and potential energies and is determined using a Boltzmann-weighted MC criterion.

$$P(1 \leftrightarrow 2) = \min\left(1, \exp\left[\left(\frac{1}{k_{\mathrm{B}}T_1} - \frac{1}{k_{\mathrm{B}}T_2}\right)(U_1 - U_2)\right]\right) \tag{3.5}$$

where the probability of an exchange between replicas 1 and 2, $P(1 \leftrightarrow 2)$, is a function of the reference temperature, T, and the relevant potential energies U. k_{B} is the Boltzmann constant.

If an exchange is successful, the considered replicas exchange conformations and the atomic velocities are rescaled to match the new replica temperature. The probability of an exchange decreases rapidly with an increasing temperature difference between replicas and so no exchanges are considered beyond neighbouring replicas. A simple schematic showing the basic principle of REMD is shown in Figure 3.3. Four replicas are pictured as horizontal lines, spanning from 400 to 1000 K. Over time (following the arrows, moving right in the left-hand schematic), exchanges occur between neighbouring replicas which allows all temperature replicas to explore the states associated with other temperatures. This allows lower temperature replicas to travel through more of the PES than they would in a standard MD simulation of the same length, which is shown simply in the right-hand plot.

The range and intervals between replica temperatures is an important simulation design consideration. Typically the lowest temperature represents the desired replica for sampling and the highest temperature is selected in order to ensure energy barriers present in the PES will be crossed. This usually means the temperature range spans across several phases. The number of replicas and their corresponding temperatures are often selected follow an exponential distribution in order to keep the overlap of potential energy distributions, and

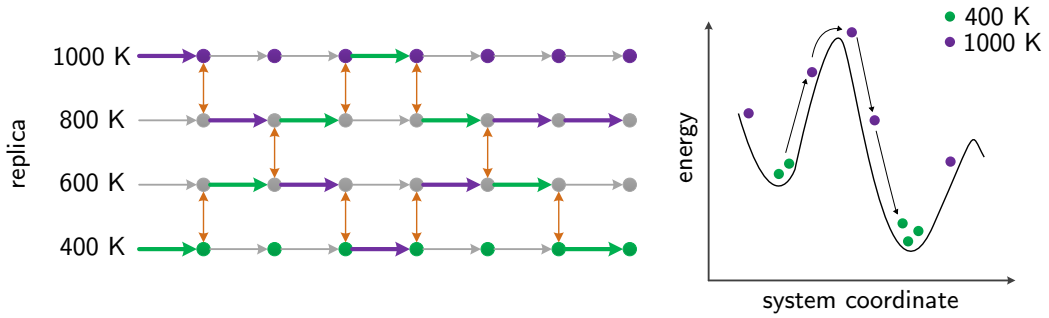


Fig. 3.3 Schematic showing exchanges between four temperature replicas over time in an REMD simulation. Four replicas at 400 (green), 600, 800, and 1000 (purple) K are represented as horizontal lines on the left, changing over time as the arrows move right. The orange arrows show when an exchange between replicas is accepted and the atomic coordinates of the neighbouring replicas are swapped. This prevents lower temperature replicas from being trapped in local minima and allows them to explore more of the PES, as illustrated in the right hand plot.

thus the resulting exchange rate, constant across all replicas. This ensures that the exchanges between replicas have meaning since neighbouring replica states have significant overlap, as long as the energy landscapes are similar across temperature space. If exchanging replicas produces valid ensembles, the probability of successful exchange attempts is high enough, and good mixing of replicas over ensembles is observed, then the resulting sampling is statistically correct and near to the ergodic ideal. Replica mixing efficiency and exchange acceptance probability are tools to analyse the effectiveness of an REMD simulation. Empirical and theoretical studies of optimal temperature distributions across replicas found that an exchange acceptance probability around 0.20 provides the best performance [268, 190].

The most significant challenge with REMD is that it is computationally expensive, due to the need to simulate many independent copies of the system of interest. This is because simulating M replicas of the system of interest requires on the order of M times more computational effort than one replica. However, as long as the temperature ranges are chosen carefully, REMD simulations are more efficient than a standard, single-temperature simulation [244]. In particular, the low temperature replicas within REMD are able to sample regions of the phase space that classic MD would not have been able to access even in a simulation M times as long. In this way, REMD is most beneficial if significant parallel computing capacity is available, although it is still not feasible to examine large systems or large temperature ranges. Despite its weaknesses, REMD is a popular method

to improve sampling efficiency, and is used in physics, chemistry, biology, engineering and materials science. For example, REMD provides valuable insight into the conformational behaviour of intrinsically disordered proteins [239, 241]. REMD has not been used to study PAH clusters but has been shown to be useful in determining the size-dependent melting behaviour of metallic nanoparticles, especially since it able to avoid the superheating and undercooling phenomena often associated with simulating these phase transitions [299]. Further information about theory, set up, and analysis of REMD simulations are provided in Chapters 4, 5, and 8.

3.2.5 Additional advanced computational methods

There have been significant advances in molecular modelling, including the development of a number of advanced methods that are useful for studying the self-assembly and structure of soot particles. Reactive molecular dynamics is one notable example. Built upon the foundation of classical molecular dynamics, reactive MD provides the ability to model bond breaking and forming and thus provides a bridge between empirical force fields possessing rigid connectivity and quantum mechanical methods that are computationally confined to much smaller systems. The modelling of reactive events is typically conducted using bond order dependency calculations and supports nanosecond dynamics for systems up to 10^7 atoms. Non-reactive force fields enforce rigid connectivity and thus typically only calculate Coulomb and van der Waals forces between atom pairs that do not share a bond or valence angle. In contrast, reactive force fields calculate these non-bonded forces between all bond pairs regardless of connectivity. Due to the fact that atomic charges and bond orders are able to change at each time step in a reactive MD simulation, this method requires a time step that is significantly smaller than that used in classical MD, often around 0.1 fs. Over the two decades, reactive molecular dynamics simulations have been increasingly used in a variety of novel research applications where reactions play an important role [296], including combustion systems [354, 214].

Many advanced molecular modelling techniques seek to sample the phase space more efficiently or extract more information out of simulation results. The former often involves enhanced sampling methods that speed up sampling using high temperatures, such as REMD and parallel tempering MC, or a biasing potential, such as umbrella sampling. There also exist a number of methods to extract relevant information from multi-ensemble molecular modelling simulations. One such method is the weighted histogram analysis method [192, 108], which allows expectations from the canonical ensemble at any temperature of interest to be computed from configurations generated by individual canonical simulations at differ-

ent temperatures. This method provides a powerful tool to extract density of states and thermodynamic properties from converged multi-ensemble simulations [70].

3.2.6 Limitations of molecular modelling

Molecular modelling techniques provide insight into small scale atomic interactions, however as with everything, these methods have limitations. It is important to understand and be aware of these limitations when using and analysing molecular simulations.

One obvious assumption in MD is that the classical equations of motion are used to describe atom movement instead of quantum methods. In addition, force fields are described as a function of atomic positions only and therefore do not take into account the motion of electrons. This is called the Born-Oppenheimer approximation and it assumes that electrons adjust their dynamics instantly when atomic positions change. This assumption is rooted in the fact that nuclei are much heavier than electrons and provides a good approximation for many systems studied. Electrons are always considered to be in their ground state, which is especially important when considering chemical reactions. Thus properties that depend on the electronic distribution in a molecule cannot be generated from MD simulations. In addition, force fields use empirical parameters and therefore their accuracy is highly dependent on the systems used for parametrisation. This means that an understanding of parameter assumptions and applicability is required when using a force field not designed for the system being investigated. Force fields are pair-additive, which means that all non-bonded forces are approximated as the sum of non-bonded pair interactions. The use of cut-offs for long-range interactions, such as Coulomb interactions, may cause unphysical behaviour at boundary points. Simulation boundary conditions also impose artificial effects that may be significant if not addressed adequately. In addition, most force fields are unable to properly describe chemical reactions due to the aforementioned assumptions regarding electrons.

MC minimisation simulations are preferred for systems in which large energy barriers may cause poor conformational sampling in dynamic methods. This is because the random jumps in a MC simulation allow the sampling to easily cross barriers, since it is only dependent on the energy before and after the jump. In addition, unlike MD simulations, MC minimisation methods do not evaluate forces to determine incremental atomic motions, but instead impose relatively large motions on the system and determine whether or not the new configuration is energetically feasible. Since the MC simulation jumps from conformation to conformation, it does not evolve smoothly through time and thus cannot provide time-dependent quantities. In contrast, MD looks directly at the dynamic behaviour of a system through the evaluation

of its movement as a function of time. MD simulations are often preferable for liquid simulations, when the energy exchange caused by molecular collisions allows energy barriers to be crossed. In these cases, MC simulations are often inefficient since there is a high probability of a random move resulting in a rejected new configuration in which molecules physically overlap. MD and MC minimisation methods require unique considerations and provide different information. In modelling soot particle stability and behaviour, it is of interest to evaluate both the temporal behaviour of the clusters at different temperatures and the minimum energy configurations and thus a combination of MD and MC minimisation simulations are beneficial.

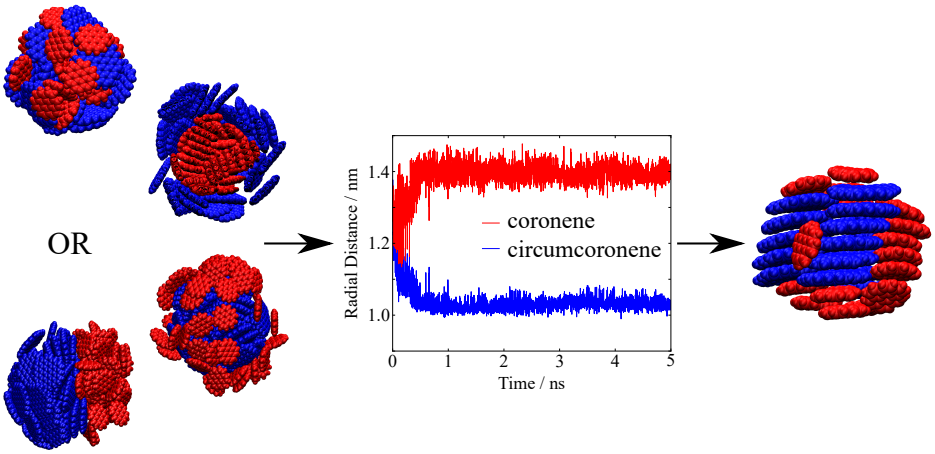
Chapter 4

Structure of particles containing heterogeneous PAHs

The morphologies of heterogeneous clusters of polycyclic aromatic hydrocarbons (PAHs) are investigated using molecular modelling. Clusters of up to 100 molecules containing combinations of the different sized PAHs circumcoronene, coronene, ovalene, or pyrene are evaluated. Replica exchange molecular dynamics simulations using an all-atom force field parametrised for PAHs sample many configurations at high and low temperatures to determine stable low energy structures. The resulting cluster structures are evaluated using molecular radial distances and coordination numbers, and are found to be independent of initial configuration and the cluster sizes studied. Stable clusters consist of stacked PAHs in a core-shell structure, where the larger PAHs are found closer to the cluster core and the smaller PAHs are located on the cluster surface. This chapter provides novel insight into the molecular partitioning of heterogeneous aromatic clusters, with particular relevance to the structure of nascent soot particles.

Highlights:

- The morphology of heterogeneous PAH clusters are investigated using replica exchange molecular dynamics.
- Stable heterogeneous PAH clusters consist of stacked PAHs in a core-shell structure, with larger PAHs in the core and smaller PAHs forming the shell.
- These low energy cluster structures are independent of initial configurations.



The work presented in this chapter was published as a paper in Carbon by Bowal et al. [43].

4.1 Introduction

In order to describe the behaviour and nature of soot particles it is necessary to understand the interactions of their components. Electronic structure calculations, such as symmetry adapted perturbation theory (SAPT(DFT)) and coupled cluster theory (CCSD(T)), provide accurate intermolecular binding energies and configurations of aromatic dimers. For small aromatics such as benzene, the T-shaped configuration (in which aromatic planes are perpendicular to each other) is the most stable, followed closely by the parallel-displaced configuration (aromatic planes are staggered face-to-face), while the sandwich configuration (aligned face-to-face) is the least favoured [157, 197, 304, 258]. This is because the hydrogen termination and π-aromatic network of these molecules creates a strong quadrupolar charge distribution that electrostatically stabilises the T-shaped configuration [330]. In contrast, the interactions between larger PAH dimers are dominated by dispersion interactions and the T-shaped configuration is substantially less stable than the parallel configurations [329, 128, 257].

Modelling aromatic clusters containing more than two molecules requires efficient parametrisation of the intermolecular interactions in order to achieve reasonable computational cost. Early approaches used the 12-6 Lennard-Jones power law functional form and atom-centred point charges fitted to the electrostatic potential [267, 149]. Improvements were made, first with a steeper exponential repulsion term [360], and then with high accuracy SAPT(DFT) calculations to generate an accurate force field, isoPAHAP [324, 327]. This PAH potential has been shown to be a transferable and highly accurate potential for systems containing planar pericondensed PAHs, for example reproducing SAPT(DFT) results of coronene dimer interaction energies, the second virial coefficients of benzene, and theoretical and experimental values of the graphite exfoliation energy [326]. Cluster stability and morphology are sensitive to the energies and potential functional form and therefore specific potentials, such as isoPAHAP, are better able to capture PAH system behaviour compared to generalised hydrocarbon potentials [131, 252, 267].

The dynamics and morphologies of homogeneous PAH clusters, containing a single PAH species, have been explored using molecular dynamics (MD) and global optimisation methods. This has allowed cluster properties such as melting points [64, 61, 37], surface accessibility [60], and gas interactions [132, 63] to be determined for PAH clusters. The morphologies of homogeneous PAH clusters have been found to be highly dependent on the type and number of PAHs present [317, 22, 267]. Small PAH clusters (for example, containing 2–6 coronene molecules) favour the formation of a single stack. Once the size of the cluster is approximately equal to the diameter of the PAH monomer (corresponding to 6 molecules for a

coronene cluster), the stacks split into two and rearrange to form twisted stacks at 90° angles, known as a handshake motif [152, 267, 22, 153]. Larger clusters (for example, 12+ coronene molecules) adopt parallel stacks in a herringbone-like motif which match the bulk crystal structures [184]. These stacked morphologies, predicted by global optimisation methods and MD studies using the isoPAHAP force field [267, 327, 252], have been confirmed recently using scanning tunnelling microscopy of coronene clusters [351].

Homogeneous clusters are often used as simple analogues of soot particles and interstellar dust even though these particles are known to contain a distribution of PAH molecule sizes, chiefly due to the increased complexity and computational cost of simulating systems containing more than one type of PAH. Only a few studies, by Wales *et al.* [267, 153], have been conducted examining heterogeneous PAH clusters. Global optimisations of heterogeneous clusters with binary combinations of coronene, hexabenzocoronene, or circumcoronene were performed using basin-hopping methods, first with an all-atom Lennard-Jones potential of clusters containing up to 10 PAHs [267] and then extended to larger clusters of up to 20 PAHs using a coarse-grained method that treats each molecule as a flattened ellipsoid [153]. In both studies, stable heterogeneous clusters showed similar structural motifs to those seen in minimised homogeneous PAH clusters, changing from single stack, handshake, and herringbone-like motifs as the number of PAHs increased. Within these stacked configurations, however, the two molecule types behaved differently, with the interactions between larger PAH interactions dominating and the smaller PAHs often capping the ends of the stacks. The coarse-grained results suggest that heterogeneous PAH clusters form stable structures in which the two molecule types are separated, although the morphology of this separation (manifesting in a core-shell or asymmetric configuration, for example) is unknown due to the fact that the clusters studied were not large enough to form more than two stacks.

Partitioning has been observed experimentally in mature soot particles, with high resolution transmission electron microscopy images showing a disordered particle core containing small fringes surrounded by larger stacked fringes [168, 10, 224, 40, 265, 16, 187, 249]. This partitioning may be physical in nature, caused by intermolecular interactions of a liquid-like particle at high temperatures in a flame. Chemical reactions, such as the carbonisation of the particle shell to form larger cross-linked molecules, may also be the source of this nanostructure. The simulation of large soot sized clusters would provide information into the contribution of intermolecular interactions in the formation of a partitioned core-shell nanostructure.

The aim of this chapter is to investigate the morphologies of heterogeneous PAH clusters, with a focus on understanding whether the core-shell structure of soot is due to physical

partitioning of different sized PAHs. This is done by investigating the internal structure of heterogeneous clusters containing up to 100 PAHs to understand molecule partitioning at a nanoparticle size relevant to nascent soot. Replica exchange molecular dynamics allows molecular rearrangements and provides stable configurations of clusters consisting of binary and non-binary combinations of circumcoronene, coronene, ovalene, and pyrene molecules. The movement and stable orientation of molecules is evaluated qualitatively, as well as by the calculation of average radial distances and molecular coordination numbers, and related to experimentally observed soot particle nanostructures. The molecular interactions involved in stable heterogeneous structures provide insight into molecule size distributions and nanoparticle morphology not found in homogeneous PAH clusters.

4.2 Computational methods

Heterogeneous PAH clusters, considered here to be clusters containing two or more types of PAHs, are examined. The following molecule combinations are considered: clusters containing coronene (COR) and circumcoronene (CIR), clusters containing pyrene (PYR) and ovalene (OVA), and clusters containing all four of the above PAHs. These molecules are selected because they span the range of PAH sizes observed in nascent soot particles, from 4–19 rings (0.5–1.1 nm) [39]. Clusters containing 32 molecules in equal molecular ratios are first assessed using several initial configurations. Large binary and non-binary clusters containing 100 molecules are then studied. In all cases, the notation AAA_xBBB_x will be used to describe the clusters studied, where AAA and BBB refer to the molecule types and x gives the number of corresponding molecules in the cluster. For example, $CIR_{50}COR_{50}$ indicates a cluster containing 100 molecules, made up of 50 circumcoronene and 50 coronene molecules.

Low energy systems indicate likely PAH cluster configurations which could be representative of soot nanoparticles sampled from a flame. In this chapter, REMD simulations are used to rapidly determine low energy configurations of PAH clusters. The simulations are initialised in four independent cluster configurations, defined by how the two molecule types are situated: mixed, janus, and two core-shell orientations (shown for a $CIR_{50}COR_{50}$ cluster in Figure 4.1(a-d)). All four initial configurations are simulated for the 32 molecule clusters, while only the mixed initial configuration is examined for the 100 molecule clusters. Clusters are created using the PACKMOL software [222] with the cluster radii selected to generate clusters of approximately 1.2 g/cm^3, corresponding to predicted soot densities [323]. Excess

energy is removed by an energy minimisation step using the steepest descent algorithm, followed by the low-memory Broyden-Fletcher-Goldfarb-Shanno method [48].

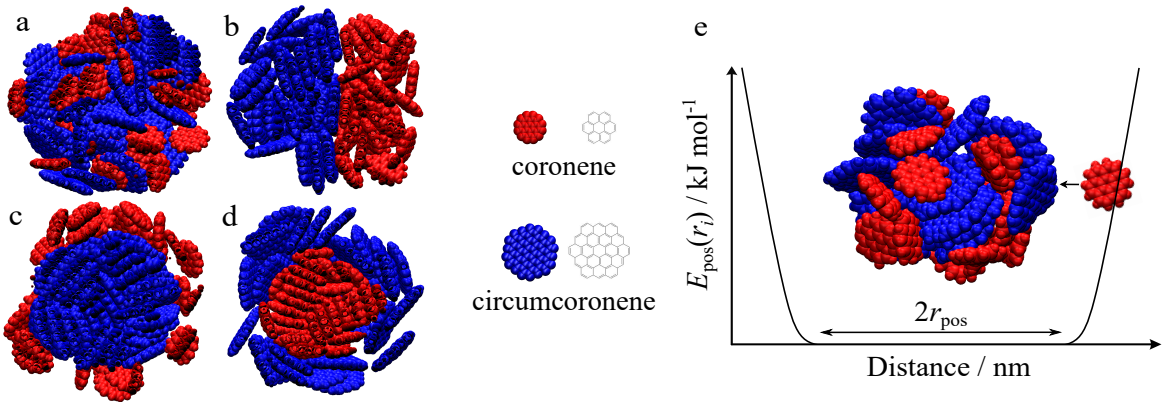


Fig. 4.1 Initial cluster geometries are (a) mixed, (b) janus, (c) circumcoronene-core coronene-shell, and (d) coronene-core circumcoronene-shell cluster configurations, shown in cross-section. Circumcoronene molecules are shown as blue, coronene molecules are shown as red. (e) Schematic showing the use of a flat-bottomed position potential to maintain a constant cluster size. The potential does not act on molecules within the cluster, only on those that leave the cluster radius, r_{pos}. $CIR_{50}COR_{50}$ clusters are shown here and all other clusters are set up in the same way.

The canonical ensemble, defined by maintaining a constant number of atoms, system volume, and temperature, is sampled using chain of 10 Nosé-Hoover thermostats for temperature control. A velocity Verlet integrator [336] is used to advance the configuration in time and all simulations are conducted *in vacuo* without periodic boundary conditions. Intramolecular forces are determined using the OPLS-AA force field [179] for molecular bonds, angles, dihedral and improper dihedral angles. This potential has been parametrised for liquid systems and used successfully in previous studies of PAH clusters [327, 71, 64, 102, 63]. The isoPAHAP intermolecular potential [324] is used to describe interactions between PAHs and intermolecular cut-offs are set to 3.0 nm.

A 5 ns trajectory of each system is simulated using a time step of 1 fs. Replica exchange attempts are made every 100 fs, since frequent exchanges are shown to increase efficiency without affecting the ensemble being sampled [301, 302]. Temperature distributions are $400-1500$ K for the CIR_xCOR_x clusters, $200-800$ K for the OVA_xPYR_x clusters, and $200-1500$ K for the $CIR_xOVA_xCOR_xPYR_x$ clusters, selected so that replicas span the temperatures of solid and liquid phases. Replica temperatures were selected using an exponential temperature distribution

$$T_j = T_0 \exp(mj) \tag{4.1}$$

where T_j refers to the temperature (in K) at replica j, T_0 is the temperature (K) at replica 0, and m is a parameter which achieves the desired temperature range. For the 32 molecule systems studied, the replica index j spans from $0-39$ and $m = 0.035$ and for the binary 100 molecule systems $j = 0-59$ and $m = 0.024$. Forty replicas are used for the 32 molecule clusters and 60 replicas for the 100 molecule clusters in order to provide an overlap in potential energies of neighbouring replicas that ensures a reasonably consistent exchange rate and efficient mixing.

In order to efficiently accept attempted replica exchanges, there must be adequate overlap of potential energy distributions between neighbouring temperature replicas [189]. Figure 4.2 shows the potential energy distribution of ten neighbouring replicas in a $CIR_{16}COR_{16}$ REMD simulation. The energy curves are broader for higher temperature replicas and it is clear that there is significant overlap between accessible energies of adjacent replica pairs, indicating good exchange rate probabilities between neighbouring replicas for this system.

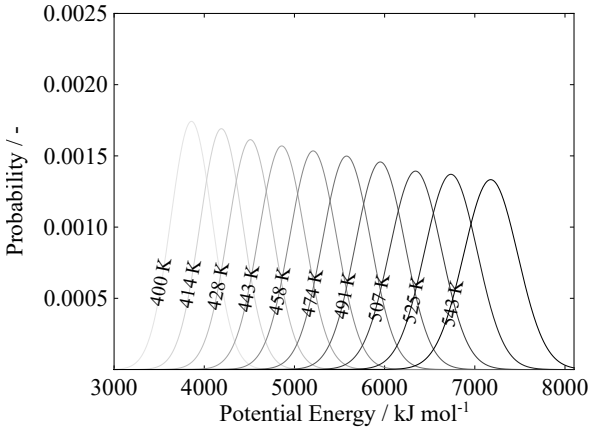


Fig. 4.2 Potential energy distributions of neighbouring replicas from a 5 ns REMD simulation of $CIR_{16}COR_{16}$. Only the 10 lowest temperature replicas, corresponding to $400-543$ K, are shown, but higher temperature replicas show the same trends.

The effectiveness of an REMD simulation relies on the proper exchange of states between replicas so that the low temperature states are able to sample the high temperature configurations and vice versa. Replica exchange acceptance is a good indication of the movement between replicas and it is found empirically and theoretically that an exchange acceptance of approximately 0.2 provided the best accuracy for a given computational time [268, 190]. Exchange acceptances for simulations conducted in this chapter are between 0.22 and 0.43, which indicates a good balance between equilibration within replicas and exchange between replicas. However the exchange acceptance is not a complete metric of REMD effectiveness since repeated exchanges between the same two replicas is treated in the same way as sequen-

tial exchanges between the full range of replicas. To ensure replicas are sampling the entire conformational space available, replica mixing is assessed by observing the movement of replica ensembles across the replica/temperature space. For all REMD simulations conducted, the replica ensembles show good sampling across replica conformations and temperatures, which is shown in a temperature trajectory for one replica ensemble in Figure 4.3. This shows the movement of a replica ensemble across the configurational space corresponding to different system temperatures. All minimisation and MD simulations are conducted using GROMACS 5.1.4 [4]. Purpose-made scripts are used to the process the output and VMD [162] provides visualisations.

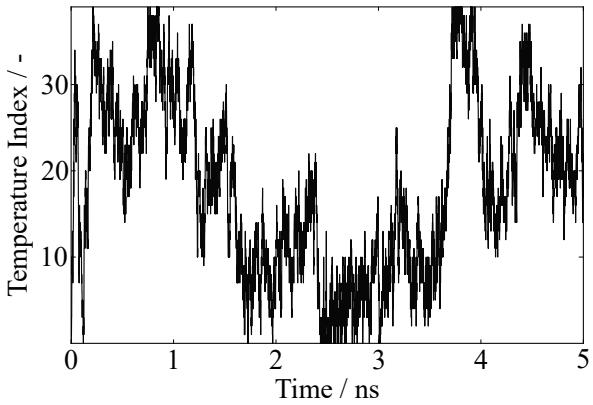


Fig. 4.3 Temperature trajectory of the lowest energy replica for a $CIR_{16}COR_{16}$ cluster. This shows the movement of a replica ensemble across the configurational space corresponding to different system temperatures. Similar configurational movement is seen for all replicas within each REMD system simulated.

Due to the size difference between molecule types, evaporation of the smaller PAHs from the cluster is observed at temperatures below the melting point of the larger PAHs. To access higher energy configurations without loss of molecules from the cluster, a position potential is implemented. This added potential acts on evaporated molecules to restrain molecules within a spherical volume. As shown in Figure 4.1(e), this flat-bottomed potential only applies a force on a molecule once it leaves the cluster, allowing all molecules within the cluster to move unaffected by the potential. The force applied acts towards the centre of the spherical volume and is independent of PAH size. The spherical position potential is as shown in Equation 4.2.

$$E_{\mathrm{pos}}(r_i) = \frac{1}{2} k_{\mathrm{pos}} \left(|r_i - R| - r_{\mathrm{pos}} \right)^2 \mathrm{H} \left(|r_i - R| - r_{\mathrm{pos}} \right) \tag{4.2}$$

where E_{pos} is the position potential energy for a single atom i in kJ mol^{-1}, r_i is the coordinate position of atom i, R is the geometric centre of the cluster, r_{pos} is the position potential radius, $k_{pos} = 1000$ kJ mol^{-1} nm^{-2} is the force constant, and H is the Heavyside step function. The position potential is necessary for minimum energy configurations to be attained since it allows the replicas to efficiently sample configurational space. Importantly, this applied potential is negligible at temperatures below the cluster melting point and low energy configurations are not affected by the position potential, which is shown in the following results.

For all clusters examined, molecular configurations are assessed by evaluating the average distance, r, of each molecule type from the geometric centre of the cluster. Additionally, molecular coordination numbers, CN, are calculated as the cumulative radial distribution function using molecule geometric centres, averaged over each molecule type. This provides an indication of the number of neighbours for each molecule type within the cluster. The maximum radial distance used in all calculations is 0.5 nm, selected to encompass the nearest interacting molecules only and provide CN values from zero to two, with $CN = 0$ corresponding to an isolated molecule, $CN = 1$ corresponding to a molecule interacting with a single other molecule, and $CN = 2$ corresponding to a molecule sandwiched between two others.

4.3 Results

Before considering the low energy configurations provided by REMD, standard MD simulations are conducted to evaluate the impact of the position potential. These simulations allow the dynamics of individual temperature systems to be assessed and provide a confirmation that sufficient configurational sampling is achieved for the REMD simulations. In particular, it is important to assess whether high energy replicas are able to rigorously sample the configuration space and low energy replicas are unaffected by the position potential. Clusters are initialised in the four different configurations discussed previously and simulated at a high temperature (1600 K for CIR_xCOR_x clusters and 800 K for OVA_xPYR_x clusters) for 5 ns.

Figure 4.4(a)-(d) shows initial and final snapshots for the four starting configurations of $CIR_{50}COR_{50}$ clusters. It can be seen that regardless of initial configuration, the final clusters exhibit a well mixed configuration, which is expected of liquid clusters for which the kinetic energy of the molecules is higher than their intermolecular energy. This indicates that these high energy replicas are able to provide the configurational sampling required for REMD

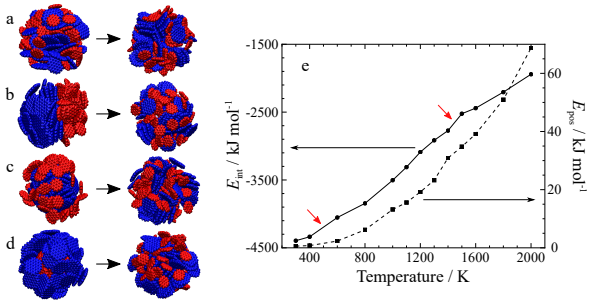


Fig. 4.4 Initial and final configurations of $CIR_{50}COR_{50}$ clusters from high energy MD simulations. As in Figure 4.1, initial geometries are (a) mixed, (b) janus, and (c), (d) core-shell cluster configurations. Larger molecules (CIR, OVA) are coloured blue, and smaller molecules (COR, PYR) are red. (e) shows the energies of a $CIR_{16}COR_{16}$ cluster across a range of temperatures. E_{int} (circles) indicates the cluster intermolecular energy and E_{pos} (squares) indicates the energy contribution of the applied position potential. Red arrows indicate approximate melting point locations of the two cluster components. All energies shown are the average cluster energies for the second half of the simulation.

due to the stochastic nature of molecule movement. These results are important because they show that at high energies, the position potential keeps the cluster intact but allows mixing to occur so that a particular structured morphology is not favoured due to the potential.

Figure 4.4(e) shows the average intermolecular energy (E_{int}) and position potential energy (E_{pos}) over a wide range of temperatures for a $CIR_{16}COR_{16}$ cluster. The sharper slopes of the change in the intermolecular energy between 1400–1500 K and 400–600 K (shown by red arrows) indicate the approximate melting point locations of the circumcoronene and coronene components, respectively. These are shifted to higher temperatures than expected based on results from analogous homogeneous clusters [64, 61], due to a pseudo-pressure effect caused by the external forces applied on the cluster by the position potential. Importantly, however, the energy contribution of the position potential is negligible at low energy replicas. This is crucial in allowing stable cluster configurations to be determined by the low energy simulations without influence from the position potential. These results show that REMD can use a position potential to sample the cluster configuration space at higher energy states and provide accurate low energy configurations that are not affected by the position potential.

REMD is applied to clusters initialised in four different configurations and the molecule size separation within the resulting structures is examined. Larger clusters containing 100 molecules, corresponding to the size of young soot particles, are then simulated. The resulting clusters are evaluated both qualitatively and quantitatively in order to provide an understanding of the interactions and structures of heterogeneous clusters.

Figure 4.5 shows the movement of the two molecule types over the duration of the REMD simulations for $CIR_{16}COR_{16}$ clusters, measured as the radial distance from the cluster centre over time. The low energy replicas show a distinct separation of the molecule types, with the larger molecules residing closer to the cluster centre and the smaller molecules situated farther away. This molecule type partitioning is not dependent on the initial cluster configurations, shown by dashed lines. The high energy replicas, shown as insets, do not show the same molecule type partitioning behaviour over time. Instead, the significant overlap between radial distance values of both molecule types throughout the simulation suggests good mixing for all molecules. The high energy replicas possess higher entropy and have radial distance values with a larger spread than those seen in the low energy replicas, indicating that the molecules are much more mobile at high energy. Again, the behaviour of these replicas is independent of initial cluster configuration.

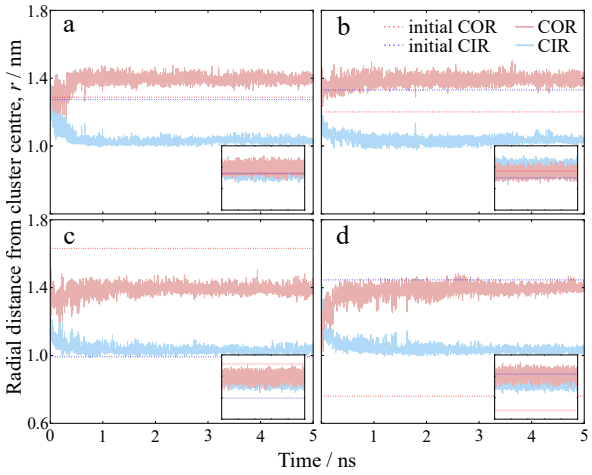


Fig. 4.5 Radial distance of each molecule type over time for $CIR_{16}COR_{16}$ clusters at the lowest energy replica, initialised with different configurations: (a) mixed, (b) janus, (c) CIR-core COR-shell, and (d) COR-core CIR-shell. Dashed lines indicate the initial configuration radii. The inset figures show the corresponding high temperature replicas using the same axes scales. The behaviour in larger CIR_xCOR_x clusters, all OVA_xPYR_x clusters, and $CIR_xOVA_xCOR_xPYR_x$ clusters follow the same trends.

Initial and final cluster morphologies are shown in Figure 4.6. These cluster snapshots show clearly that the arrangement of molecule types in the final cluster configurations are similar, regardless of initial geometry or molecule types. In addition, the stable structures (shown on the right of the arrows) present a stacked morphology, indicating that dispersion interactions dominate. This stacked structure is expected due to the highly anisotropic nature of the molecule interaction strengths [104] and has been observed experimentally in PAH clusters [351]. As predicted by quantum calculations of PAH dimers, the slipped-parallel and sandwich molecule arrangements are dominant. As observed in previous work on PAH

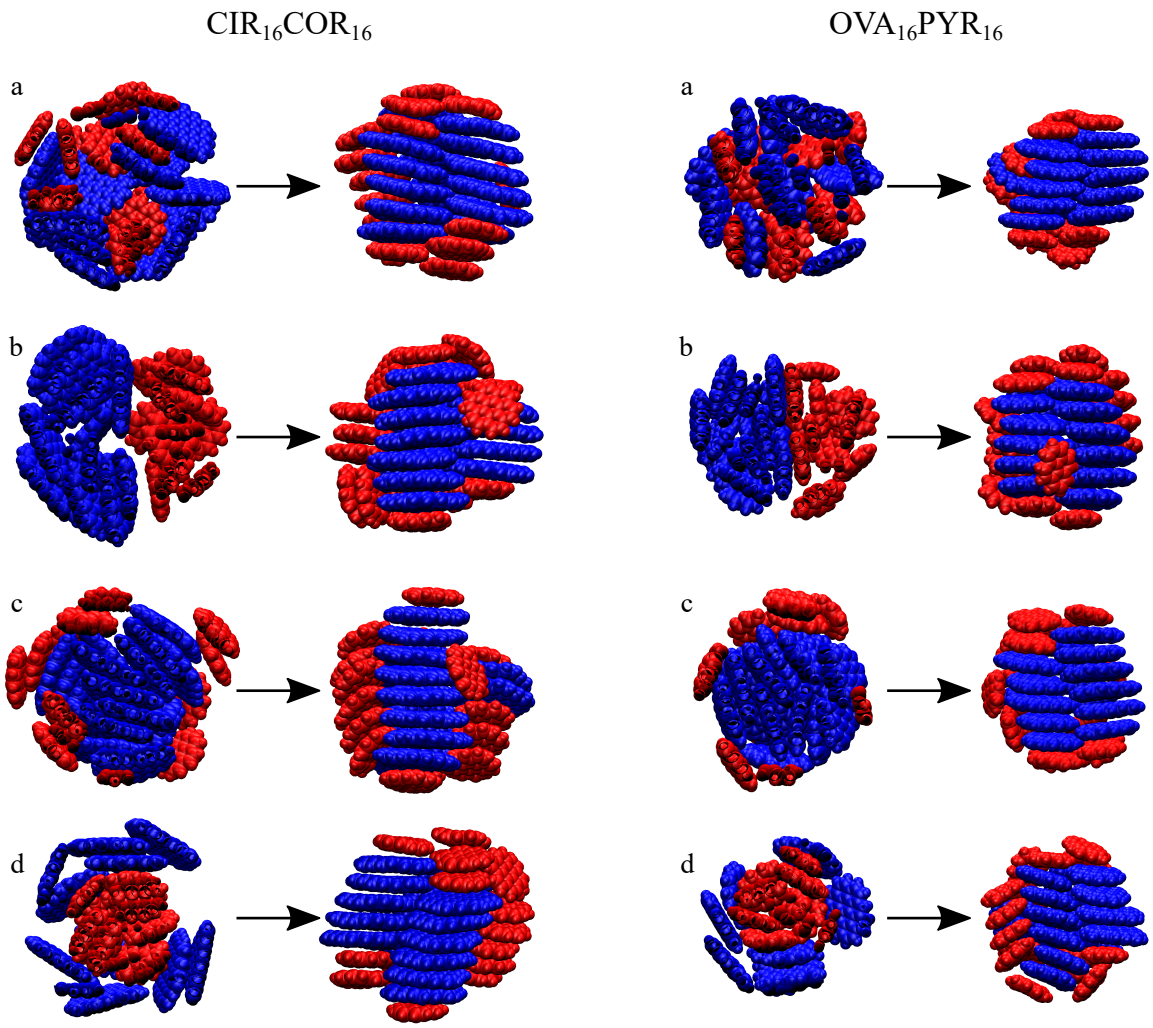


Fig. 4.6 Initial and final configurations of $CIR_{16}COR_{16}$ and $OVA_{16}PYR_{16}$ clusters from the lowest energy replicas. The final stable configurations are similar regardless of initial configuration. As in Figure 4.1, initial geometries are (a) mixed, (b) janus, and (c), (d) core-shell cluster configurations, shown in cross-section. Larger molecules (CIR, OVA) are coloured blue, and smaller molecules (COR, PYR) are red.

clusters [153], the slipped-parallel interactions are not dominant enough to produce the helical morphologies seen in theoretical discoid studies [105, 58]. Parallel stacks show the tilted behaviour characteristic of a herringbone-like motif, which is the motif found in bulk crystal structures [184]. Some T-shaped interactions are seen between the two molecule types, highlighting that the electrostatic interactions of these aromatics do contribute [246].

A partitioning of the molecule types is evident, with the larger molecules forming parallel stacks in the core of the cluster and the smaller molecules capping the ends of these stacks

and interacting more weakly in outer stacks or as individual molecules. The stacks formed by the larger PAHs contain more molecules than those consisting of the smaller PAHs, although none are larger than the limit due to the dissociation energy observed in analogous homogeneous clusters [267]. This suggests that the molecules in the core behave similarly to homogeneous clusters of the same size studied previously. The interactions of the smaller molecules around the outside of these parallel stacks indicate a clear priority to maximise carbon-carbon interactions, as observed in previous studies on homogeneous clusters [153], leading to configurations in which the larger molecules are located closer to the cluster centre. This strong molecule size partitioning was not seen in previous work where only small clusters consisting of two molecule pairs were considered [267].

Cluster snapshots provide visual results of the simulations but do not provide a complete analysis. In order to further understand the molecular distributions of these clusters, the radial distributions and coordination numbers are calculated to quantify partitioning and surface composition, respectively.

Figure 4.7 shows the distributions of the radial distance, measured from the centre of cluster to the centre of each molecule, for the two molecule types for all $CIR_{16}COR_{16}$ clusters over the final 3 ns of the simulations. The average radial distances, $\langle r \rangle$, from Table 4.1 are shown as dashed vertical lines. These figures provide an indication of the overall partitioning of the two molecule sizes over time and allow the penetration of a given molecule type within the cluster centre to be quantified. All clusters show a similar structure in which the larger molecules are found at lower radial distances from the cluster centre than the smaller molecules. The spread of the histograms show movement of the molecules with time, but the two molecule types remain in distinct regions without significant overlap. This indicates that the larger molecules stay within the cluster core while smaller molecules are unable to penetrate into the cluster centre and are instead favoured in an outer shell. Table 4.1 provides the radial distances for all clusters studied and it is seen that in all cases, the larger molecules have a smaller radial distance than the smaller molecules in the cluster, indicating a core-shell structure. In addition, the difference between the molecule type radii corresponds to a single surrounding layer of smaller molecules.

Evaluation of the mixing between molecule types and the surface composition is provided by atomic distributions of the radial distances and coordination numbers for a single point in time, shown in Figure 4.9. An overall partitioning of the two molecule types is again seen, with the atoms within the smaller molecules occupying regions of larger radial distances than the atoms within the larger molecules. There is an overlap of these atomic radial distances, not seen in the equilibrium distributions, which indicates that the two molecule types do

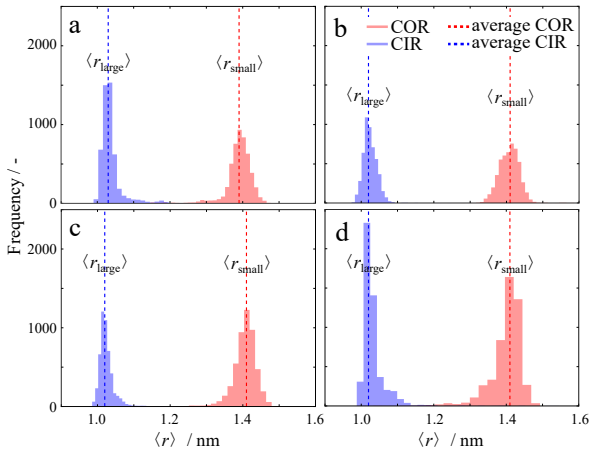


Fig. 4.7 Molecular equilibrium radial distance distributions of $CIR_{16}COR_{16}$ configurations, considering the different initial geometries: (a) mixed, (b) janus, (c) CIR-core COR-shell, and (d) COR-core CIR-shell. Dashed vertical lines correspond to average radii and show that larger CIR molecules $\langle r_{\text{large}} \rangle$ are found closer to the cluster centre than smaller COR molecules $\langle r_{\text{small}} \rangle$.

interact with each other within the clusters, although not to the extent that interdigited stacks are formed. When comparing to the normalised atomic radial distance distributions corresponding to the initial starting geometries, shown in Figure 4.8, it is apparent that all clusters show an arrangement similar to the circumcoronene-core coronene-shell initial geometry (Figure 4.8c).

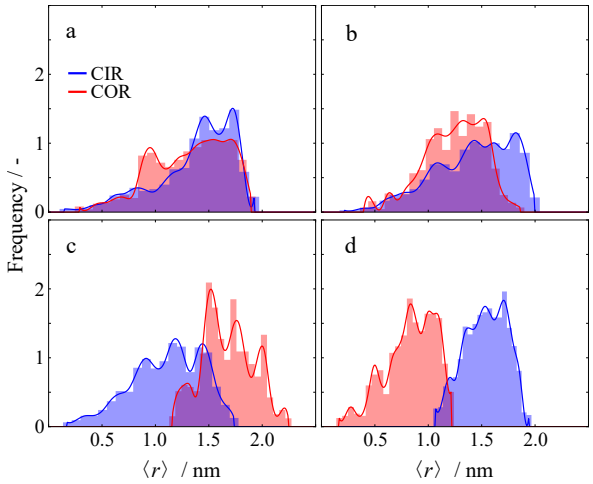


Fig. 4.8 Normalised atomic radial distance distributions of initial $CIR_{16}COR_{16}$ configurations, initialised in different configurations: (a) mixed, (b) janus, (c) CIR-core COR-shell, and (d) COR-core CIR-shell.

The shape of the peaks supports a core-shell structure since the sharp peaks corresponding to the smaller molecule atoms suggest that these molecules are found in a layer around the larger molecules. The bimodal distribution seen for both molecule types in Figure 4.9(c), with a larger and broader peak at lower radii and a small sharp peak at high radii, can be explained by considering the stacked structures formed. As seen in Figure 4.6(c) this cluster contains a longer stack in the core compared to two smaller parallel stacks found in other configurations. This produces a cluster that is slightly more elongated than the other mostly spherical clusters and thus the smaller peaks found at the larger radii values correspond to the molecules at both ends of the molecular stacks.

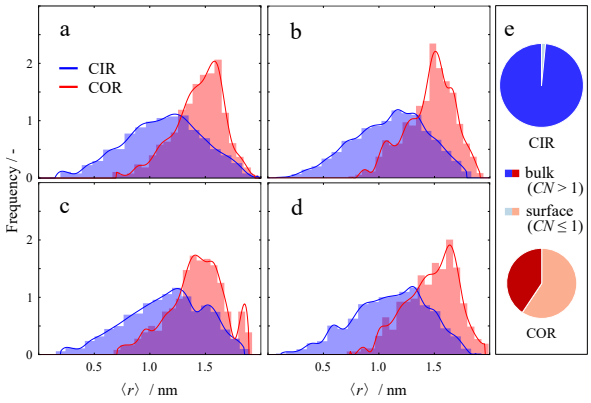


Fig. 4.9 Normalised atomic radial distance distributions of final $CIR_{16}COR_{16}$ configurations, initialised in different configurations: (a) mixed, (b) janus, (c) CIR-core COR-shell, and (d) COR-core CIR-shell. (e) displays the bulk and surface molecule distributions for all $CIR_{16}COR_{16}$ clusters, determined by coordination numbers.

The extent of stacking of each molecule type can be expressed as a coordination number (*CN*), calculated as the average number of neighbouring molecules for each molecule within a PAH type. This value is also able to provide an assignment of surface and bulk molecules in the cluster, since molecules interacting closely with only one other molecule ($CN \leq 1$) are found on the cluster surface while molecules surrounded by more than one molecule ($CN > 1$) are indicative of bulk cluster molecules. The pie charts in Figure 4.9(e) highlight the proportion of molecules within each molecule type that are situated in the bulk or surface of the clusters. The darker colours for both CIR and COR molecules refer to proportion of molecules in the bulk and lighter colours show surface molecules. Nearly all of the larger CIR molecules interact with several near neighbours and are found within the cluster, while smaller COR molecules make up most of the surface molecules and display some bulk-like characteristics due to the formation of outer stacks. Table 4.1 provides the average *CN* values for all clusters studied at their final configuration. The *CN* values for the large molecules are near two in all cases, corresponding to each molecule experiencing close interaction with

two other molecules, indicating a stacked structure. The smaller molecules have *CN* values around one, corresponding to surface molecules only located near one other molecule. Larger molecules are thus more stacked and less likely to be found on the surface of the cluster since they have higher *CN* values than the corresponding smaller molecules.

Large clusters containing 100 molecules are simulated to provide insight into soot sized nanoparticle behaviour. Final configurations of the 32 molecule clusters show that the stable configurations are not dependent on the initial structure, and thus these larger clusters are simulated only with mixed initial configurations. The resulting final configurations are shown in Figure 4.10(a) and (b). The morphologies of $CIR_{50}COR_{50}$ and $OVA_{50}PYR_{50}$ clusters are similar to those found for the smaller clusters: tilted parallel stacks, with some T-shaped interactions between molecules around the surface. Molecular stacks dominate and the average molecular radial distances $\langle r \rangle$ and *CN* values indicate that the larger molecules are found closer to the cluster centre than the smaller molecules. This suggests that unlike the changing stable morphologies found for small cluster sizes [236, 94, 267], clusters containing between 32 and 100 molecules form similar morphologies. The large cluster size allows the herringbone-like motif and partitioning between molecule sizes in a core-shell structure to be more clearly seen.

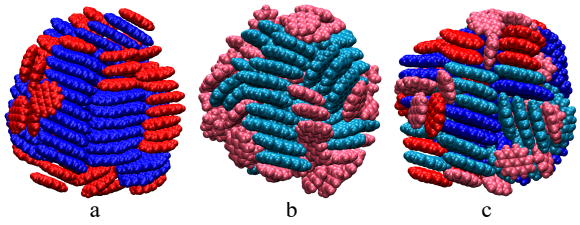


Fig. 4.10 Final configurations of large heterogeneous PAH clusters, (a) $CIR_{50}COR_{50}$, (b) $OVA_{50}PYR_{50}$, and (c) $CIR_{25}OVA_{25}COR_{25}PYR_{25}$. The colour assignments are as follows: CIR is blue, OVA is cyan, COR is red, PYR is pink.

The PAH clusters considered so far are somewhat ideal since they contain only two molecule types. This allows for rigorous modelling but perhaps could lead to results indicative of a binary cluster only. To address this, a further REMD simulation of a large non-binary PAH cluster containing 25 molecules of circumcoronene, ovalene, coronene, and pyrene (100 molecules total) was simulated, initialised in a mixed configuration. The simulation set up follows the same procedure as the previous clusters, using $j = 0-74$ and $m = 0.0275$ (within Equation 4.1 for the replica temperature distribution). The resulting cluster configuration, shown in Figure 4.10(c), shows molecule size partitioning similar to that seen in the binary clusters. The largest CIR molecules are found within the cluster core along with the large OVA molecules, while the small COR molecules and smallest PYR molecules are located

on the ends of the stacks containing the larger molecules as well as in a perpendicular arrangement around the outside of the cluster. Figure 4.11 shows the distribution of the radial distance for the four molecules within the $CIR_{25}OVA_{25}COR_{25}PYR_{25}$ cluster over the final 3 ns of the simulation. The average radial distances, $\langle r \rangle$, are shown as dashed vertical lines and show a partitioning of molecule types within the cluster. The *CN* values are also provided. These values indicate that the larger the molecule, the closer it is to the cluster core (lower $\langle r \rangle$ value) and the more highly stacked it is (higher *CN* value). In contrast to the binary clusters there is overlap between the molecule types, likely due to their similar sizes. However, a distinct core-shell structure is still observed in which the larger molecules form the core and the smaller molecules make up the shell.

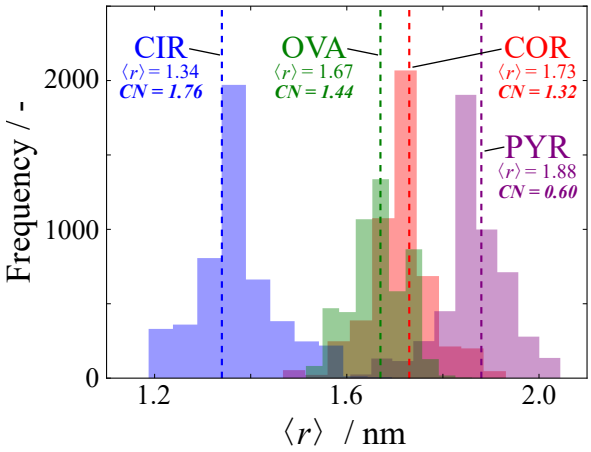


Fig. 4.11 Molecular equilibrium radial distance distributions of $CIR_{25}OVA_{25}COR_{25}PYR_{25}$ configurations, initialised in a mixed configuration. Average radii values are provided and shown as vertical dashed lines. Molecular *CN* values are also shown in bold italics.

Table 4.1 provides the binary cluster intermolecular energies, and it can be seen that this energy increases with component molecule size. The interaction energies of the large binary heterogeneous PAH clusters are between those of analogous clusters containing one type of PAH. For example, a cluster containing 50 coronene and 50 circumcoronene molecules has an approximate average intermolecular energy of $-14,000$ kJ/mol, while the energy of a cluster containing 100 coronene molecules is around $-9,400$ kJ/mol [64] and a cluster of 100 circumcoronene molecules is $-21,000$ kJ/mol [61]. Similarly, at $-13,442$ kJ/mol, the intermolecular energy of the $CIR_{25}OVA_{25}COR_{25}PYR_{25}$ cluster is approximately equal to that predicted by the mean of the four homogeneous 100 molecule cluster values.

Table 4.1 Average radial distances ($\langle r \rangle$, nm), coordination numbers (CN), and intermolecular energies ($\langle E_{\text{int}} \rangle$, kJ mol^{-1}) of heterogeneous PAH clusters from REMD simulations for different initial configurations and sizes. Subscripts refer to the large and small molecule types within each cluster. Properties are empirical equilibrium values, as indicated by angled braces, determined as the average over the final 3 ns of the simulation.

Structure	Initial configuration	$\langle r_{\text{large}} \rangle$	$\langle r_{\text{small}} \rangle$	CN_{large}	CN_{small}	$\langle E_{\text{int}} \rangle$
$CIR_{16}COR_{16}$	CIR-core COR-shell	1.02	1.41	2.00	1.56	-4590
$CIR_{16}COR_{16}$	COR-core CIR-shell	1.02	1.42	1.94	1.19	-4599
$CIR_{16}COR_{16}$	janus	1.02	1.41	2.00	0.88	-4601
$CIR_{16}COR_{16}$	mixed	1.03	1.39	2.00	1.25	-4602
$OVA_{16}PYR_{16}$	OVA-core PYR-shell	0.95	1.25	2.00	0.75	-2901
$OVA_{16}PYR_{16}$	PYR-core OVA-shell	0.95	1.19	1.94	0.94	-2901
$OVA_{16}PYR_{16}$	janus	0.87	1.21	2.00	0.88	-2907
$OVA_{16}PYR_{16}$	mixed	0.90	1.19	1.94	0.69	-2902
COR_{100} [64]	–	–	–	–	–	-9400
$CIR_{50}COR_{50}$	mixed	1.61	2.07	1.92	1.24	-14236
CIR_{100} [61]	–	–	–	–	–	-21000
PYR_{100} [64]	–	–	–	–	–	-6800
$OVA_{50}PYR_{50}$	mixed	1.40	1.73	1.78	0.82	-10244
OVA_{100} [61]	–	–	–	–	–	-12000

4.4 Discussion

The shape and interaction strength of anisotropic molecules govern the stable cluster morphologies they form [104]. In this way, the partitioning of molecule types within heterogeneous PAH clusters to form a core-shell structure may be caused by the different molecule sizes or molecule interaction strengths. It is difficult to assess their individual contributions since these two factors are coupled. Clusters of spherical Lennard-Jones (LJ) particles provide simple systems for which these components have been assessed independently and can provide insight into the PAH clusters studied here. LJ clusters made up of two different sized particles with the same interaction potential are known to configure in a core-shell structure, with larger particles on the surface and smaller particles in the core, in order to equalise surface and bulk interactions [95, 94, 276, 74]. In contrast, LJ clusters containing two particle types of the same size but different attractions partition in the opposite core-shell structure, in which the more strongly attractive particles make up the core and the more weakly attractive particles compose the shell [236]. Heterogeneous PAH clusters show a structure similar to the latter case, suggesting that the strength of intermolecular interactions plays a more dominant role in determining cluster structure than the molecular size differences.

A detailed examination of the average intermolecular interactions over the entire simulation period finds that the contribution from repulsion-dispersion dominates, with electrostatics contributing to less than 3% of the intermolecular interaction energy, for all clusters studied. The influential role of dispersion in these systems provides further understanding into the cluster structures and partitioning observed. The stacked configurations seen serve to maximise π-π interactions in contrast to electrostatics-dominated structures, which favour an edge-on structure such as a T-shaped configuration [129]. As an additive force, dispersion increases with increasing molecular weight due to the increased number of pair-wise interactions. Thus large PAHs interact strongly with each other, while interactions between a large PAH and small PAH or between two small PAHs are less energetically favoured. This leads to an arrangement in which larger PAHs prefer to occupy interior sites in order to maximise the number of nearest neighbours between carbon atoms, and smaller PAHs are located on the surface of the cluster. The heterogeneous clusters examined have a large disparity between the molecule attraction strengths due to their different sizes. The results shown, in particular the molecule type partitioning at low energies, are more pronounced than would likely be the case between PAHs with similar interaction strengths. Therefore a cluster containing PAHs with similar sizes, for example the pericondensed PAHs circumpyrene and hexabenzocoronene with 14 and 13 aromatic rings respectively, would likely exhibit more mixing at lower energies. Evidence of this increased molecular mixing is seen visually in the $CIR_{25}OVA_{25}COR_{25}PYR_{25}$ cluster configuration shown in Figure 4.10(c).

The minimum energy cluster configurations determined in this chapter indicate expected quenched soot structures, such as those obtained through experimental sampling methods in which soot is thermophoretically condensed onto a sampling grid for analysis. These heterogeneous clusters have diameters of $3-5$ nm and C/H ratios of $2-2.6$, a size and composition representative of nascent soot particles [277, 85]. In addition, stable configurations contain molecular planes perpendicular to the cluster surface, a morphology that has been identified as typical of nascent soot [187]. This is distinct from the more graphitic structure of mature soot particles where the molecular planes are parallel to the particle surface.

Cluster structures show a core-shell structure in which stacked larger molecules dominate the core of the cluster. These simulations results prompted and guided colleagues to perform the first experimental work analysing molecule size with respect to soot particle radius at different soot maturities [41]. Using high resolution transmission electron microscopy, this study showed that the molecules are slightly larger and more stacked at the core than at the surface within the youngest particles able to be sampled. This provides experimental support for the unexpected molecule size distributions of heterogeneous PAH cluster configurations determined in this chapter. The inverse morphology, where smaller fringes are found in the

particle centre and larger fringes make up the surface layers, is reported qualitatively for mature soot particles using high resolution transmission electron microscopy [40, 168, 16]. It is known that soot particles form a graphitic shell as they mature and the results in this chapter indicate that intermolecular interactions within a liquid-like particle are not responsible for the core-shell partitioning seen in mature particles. As only non-reactive molecular interactions are considered here, it is likely that chemical reactions such as carbonisation play a crucial role in determining mature soot particle structure. The identification of this inverse core-shell structure for nascent soot particles suggests their surface interactions are dominated by small PAHs. This provides a possible contributing factor for the enhanced reactivity of these young particles compared to mature soot [52].

Further work to better understand the transition from nascent to mature soot particles, likely dominated by carbonisation, and its impact on the core-shell structure is required. Curved PAH molecules have been very recently shown to form a significant component of soot particles [217] and further work (described in Chapters 7 and 8) will explore the nature of curved molecule cluster morphologies. It is expected that similar stacked motifs will be observed, although molecule type partitioning in heterogeneous clusters may be less distinct since the interaction between molecules of different sizes is enhanced for these polar systems [221]. In addition, the stericity of these curved molecules may allow heterogeneity to further stabilise clusters, for example if smaller molecules are able to pack well into the concave surface of larger molecules. While REMD is able to provide stable configurations, it comes at a very high computational cost. There is therefore still a need for efficient methods for sampling large heterogeneous aromatic clusters. This is explored further in Chapter 6.

4.5 Conclusions

The morphology of large heterogeneous PAH clusters is explored using advanced molecular modelling. Replica exchange molecular dynamics simulations across a wide temperature range are able to show that clusters initialised with different geometries produce stable configurations independent of the initial cluster structure and size. Stable clusters present a parallel stacks motif, similar to the herringbone-like motif seen in bulk PAHs. The maximisation of carbon-carbon interactions dominates the arrangements formed, leading to smaller PAHs capping the ends of otherwise homogeneous stacks of larger PAHs. This agrees with previous work looking at interactions between PAH dimers and small heterogeneous clusters. For all systems studied, it is seen that heterogeneous PAH clusters favour a core-shell structure in which the larger molecules are located closer to the cluster centre. These

results present the first simulations of soot sized heterogeneous PAH clusters and provide insight into the internal structure of non-carbonised nascent soot particles, useful for further understanding of soot formation and other similar systems.

Chapter 5

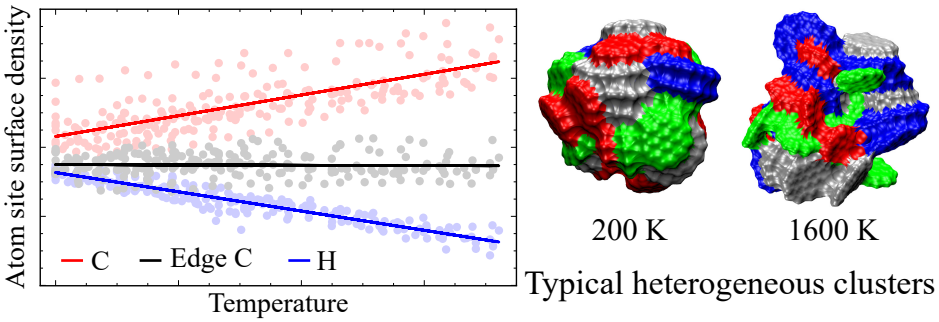
Surface properties of particles containing heterogeneous PAHs

The impact of molecular inhomogeneity on the surface properties of soot particles is investigated using replica exchange molecular dynamics and solvent-excluded surface analysis. The temperature-dependent behaviour of surface roughness and number densities of reactive sites are evaluated for particles from 1–5 nm in diameter. The percentage of carbon atoms and zig-zag sites on the particle surface are found to be independent of molecular composition, while molecule heterogeneity influences the accessible hydrogen atoms and free-edge sites. These relationships allow the prediction of surface composition for a given particle diameter. The surface densities of carbon and hydrogen atoms are explained by the morphological changes and molecule size contributions for solid-like and liquid-like configurations. Small molecules contribute significantly to the particle surface properties at low temperatures, regardless of the proportion of molecule sizes, which results in an increased density of accessible carbon atoms for heterogeneous particles. Interestingly, the surface density of edge carbon atoms and free-edge sites can be predicted from the average molecule size alone. The density of hydrogen atoms on the surface follows the average expected values from the constituent molecule sizes, suggesting that for particles containing many different molecule sizes the α parameter corresponding to the HACA mechanism converges to a linear temperature-dependent trend. This quantitative evaluation of the acces-

sibility of reactive sites for heterogeneous particles provides important information for understanding soot particle growth and oxidation.

Highlights:

- Surface composition, roughness, and site densities are explained in the context of cluster morphology.
- Surface densities of hydrogen, internal carbon, and zig-zag sites are dependent on temperature, while edge carbon and free-edge site surface densities can be predicted from average molecule size.
- Small molecules contribute more to cluster surface properties than large molecules.
- The surface area contribution from hydrogen shows a linear temperature-dependent trend with increasing molecular heterogeneity.



Typical heterogeneous clusters

The work presented in this chapter was published as a paper in the Proceedings of the Combustion Institute by Bowal et al. [45].

5.1 Introduction

The size and nature of soot particles produced in combustion processes are dependent on their mass growth and oxidation, which are governed by chemical interactions between the soot particle surface and gaseous species such as acetylene and oxygen [115, 52]. These surface reactions have been investigated in numerous experimental and theoretical studies and it is well-established that the nanostructures of soot particles, defined by the size, orientation, and organisation of constituent polycyclic aromatic hydrocarbons (PAHs), play a significant role in reactivity [237, 118]. This is due to the influence of nanostructure on the accessibility of reactive sites on the particle surface. Edge carbon (EC) atoms on the rim of PAH molecules have significantly higher reactivities (10-100 fold difference) than internal carbon (IC) atoms on the basal plane [11]. Smaller PAHs possess higher proportions of EC atoms and lower C-to-H ratios, leading to an inverse relationship between molecule size in soot particles and the apparent oxidation rate constant [365].

To capture the dynamic reactivity of soot particles in combustion processes the description of surface reactions involves an empirical parameter α, which captures the fraction of the soot surface available for surface reactions. The parameter α was first introduced by Frenklach and Wang to reproduce soot volume fractions for both high and low temperature flames [120]. The parameter α is often understood to capture the changing morphology of soot particles as a function of temperature and particle size, in particular the presence of reactive graphitic edges or unreactive basal aromatic planes on the particle surface [17]. Due to the difficulty of probing atomic properties with experimental methods, the parameter α is selected to provide the best fit between model predictions and experimental results. As a consequence, this steric parameter is not derived from soot nanostructures, leading to a variety of α values across models without a clear convergence with temperature or over time [62]. Over the last three decades significant attention has been spent to develop improved formulations of the parameter α , including explicit dependencies on temperature [181], particle size [17], and age [303]. Recently, Frenklach presented a new form of α that moves away from variable empiricism and captures instantaneous surface site concentration [116].

Developing a deeper understanding of particle reactivity requires detailed investigation of soot surface properties. Recent modelling work used atomic simulations to characterise the surface properties of PAH clusters and their interactions with gaseous molecules [60, 62, 132]. This provides a numerical basis to connect molecular arrangement with the surface roughness, number density of reactive site i on the surface (χ_i), and a formulation of α that includes the effects of temperature, particle size, and chemical composition. These studies are valuable

but are limited to homogeneous particles in which there is only one constituent molecule type. Soot particles contain molecules of many sizes which are known to partition in a core-shell structure where the larger molecules make up the particle core (nascent soot) or where larger molecules dominate the outer particle shell (mature soot), as discussed in Chapter 4. These nanostructural differences are expected to influence particle surface composition but their detailed effect on χ_i values and the parameter α have yet to be determined.

In this chapter a detailed numerical evaluation of the surface properties of heterogeneous clusters containing PAHs of different sizes and ratios is performed. It is important to note that (as with all molecular models) these clusters do not capture all possible features of a soot particle, but serve as a valuable model to vigorously assess the influence of different molecule sizes and ratios on cluster surface properties, thereby revealing how heterogeneity may influence the properties of a PAH cluster. In particular, this chapter provides insight into the reactive surface of PAH clusters and allows refinement of model parameters especially important in understanding soot surface growth. The proportion of atom and site types on the particle surface is explored as a function of diameter to highlight the impact of inhomogeneity. Surface roughness and site surface densities are explored, and the contributions of different molecule sizes are highlighted.

5.2 Methodology

The particle systems investigated are made up of pyrene (PYR), coronene (COR), ovalene (OVA), and circumcoronene (CIR) molecules. These molecules were selected because they have been identified to play a role in soot formation [229], they span the molecular sizes observed experimentally in soot particles [15, 41] with some variety in oval/circular molecular shape, and importantly, data for homogeneous clusters containing these PAHs are available in the literature, allowing for a direct comparison of the impact of heterogeneity on surface properties. Binary and quaternary particles of various sizes, containing two and four molecule types respectively, are considered (listed in Table 5.1). Two types of binary clusters, containing CIR and COR or OVA and PYR, are examined. Two types of quaternary clusters are also considered, one containing a uniform distribution of molecule types and the other with a nonuniform distribution in which the larger molecules are found in higher proportions. This latter distribution is selected as a conservative case for surface property analysis, since smaller molecules are expected to contribute more significantly due to core-shell partitioning. The molecule types are shown in Figure 5.1 along with snapshots of representative quaternary and binary particles at low and high temperatures.

Table 5.1 Heterogeneous particles considered in this chapter. Systems labelled by an asterisk were taken from the simulations presented in Chapter 4.

Cluster type	**Number of molecules**	**Particle diameter (nm)**	m	j
Binary 50% OVA, 50% PYR	4	1.34	0.107	0 – 13
	32*	2.80	0.035	0 – 39
	100*	4.15	0.024	0 – 59
Binary 50% CIR, 50% COR	4	1.54	0.093	0 – 15
	32*	3.20	0.035	0 – 39
	100*	4.78	0.024	0 – 59
Quaternary (uniform)	4	1.45	0.091	0 – 23
25% CIR, 25% OVA,	32	3.02	0.036	0 – 59
25% COR, 25% PYR	100*	4.48	0.028	0 – 74
Quaternary (nonuniform)	10	2.11	0.054	0 – 39
40% CIR, 30% OVA,	24	2.87	0.036	0 – 59
20% COR, 10% PYR	100	4.70	0.024	0 – 87

Heterogeneous particles are produced following the same procedure as the previous study detailed in Chapter 4. PAH clusters are generated by randomly placing molecules in a spherical volume using the PACKMOL software [222]. A two-step minimisation process using the steepest descent and low-memory Broyden-Fletcher-Goldfarb-Shanno algorithms removes excess energy prior to the equilibration and production simulations. Replica exchange molecular dynamics (REMD) simulations, each sampling the NVT ensemble, defined by maintaining a constant number of atoms, system volume, and temperature, are performed using GROMACS 5.1.4 [4]. REMD simulations use a large number of interacting parallel simulations to rapidly determine stable structures across a large range of temperatures. A timestep of 1 fs is used and 1 ns simulations are sufficient for particles < 3 nm while larger clusters (up to 5 nm) require 2 ns simulations to produce structurally equilibrated particles with energy drifts less than that of experimental error. The temperature ranges are selected to encompass both solid- and liquid-like configurations. In these heterogeneous clusters the temperature ranges differ with the molecular composition of the cluster, since constituent molecule sizes determine the temperature at which a cluster is liquid- or solid-like. The smallest molecule present determines the lower temperature bound, with pyrene at 200 K and coronene at 400 K. Similarly the largest molecule present determines the upper temperature bound, with ovalene at $\geq$ 800 K and circumcoronene at $\geq$ 1600 K. Replica temperatures were selected using an exponential temperature distribution as described in Chapter 4. The j and m parameters, corresponding to Equation 4.1, were selected for each cluster system to allow satisfactory exchange between replicas (discussed further below) using these temperature ranges and are reported in Table 5.1. This requires between 14 and 88 REMD replicas to maintain an acceptable replica exchange acceptance with attempted exchanges every 100 fs.

Exchange acceptances for simulations are approximately $0.2 - 0.4$, which indicates a good balance between equilibration within each replica and exchange between replicas.

A chain of 10 Nosé-Hoover thermostats is used for temperature control, a velocity Verlet integrator is used to advance the configurations in time, intermolecular cut-offs are set to 3.0 nm, and simulations are conducted *in vacuo* without periodic boundary conditions. A spherical position potential is applied using a flat-bottomed potential described in Chapter 4. This potential allows the replicas to efficiently sample configurational space and, importantly, it is negligible at temperatures below the cluster melting point and low energy configurations are not affected by the position potential. Intermolecular interactions are described using the isoPAHAP potential [324], parametrised specifically for PAH interactions. The previous detailed assessment described in Chapter 4 showed that the initial cluster arrangement and the use of a position potential do not influence the molecular structure of the equilibrated clusters sampled.

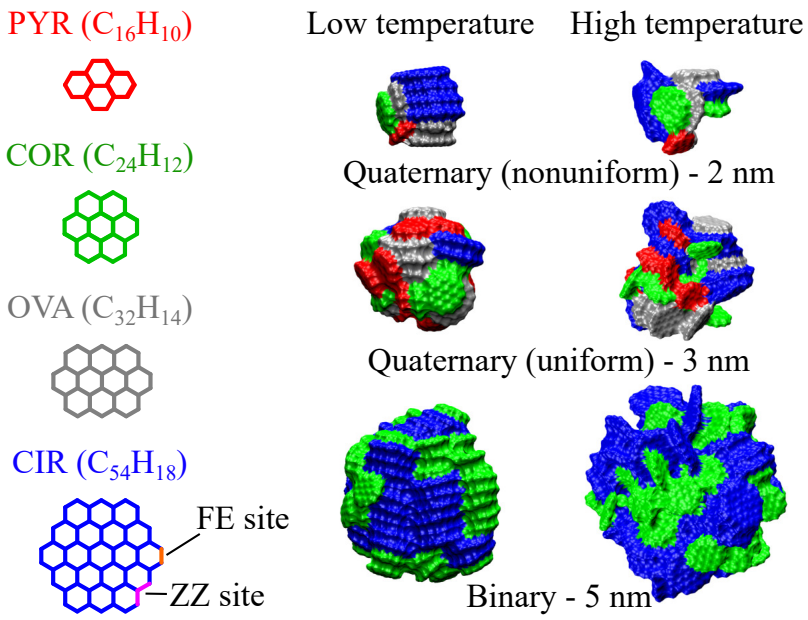


Fig. 5.1 Pyrene (red), coronene (green), ovalene (grey), and circumcoronene (blue) molecules are considered. Example free-edge and zig-zag sites are highlighted. Snapshots of representative heterogeneous particles are shown at low and high temperatures (solid-like and liquid-like configurations, respectively).

The detailed surface properties of the simulated particles are determined using solvent-excluded surface (SES) analysis, which uses a 'rolling ball' algorithm to describe the three-dimensional surface. This provides the surface area and quantitative contributions from reactive sites of interest: hydrogen (H) and carbon (C) atoms, the latter of which can be

split into IC atoms and EC atoms. EC atoms can be further defined as free-edge (FE) sites containing two carbon atoms or zig-zag (ZZ) sites made up of three carbon atoms, shown in Figure 5.1. Statistics for the detailed surface properties are collected from 6 equilibrated configurations, each extracted at 20 temperatures. Following a previous work on homogeneous particles [62], the SES analysis uses the MSMS 6.2.1 program [284] with a probe radius of 2.0 Å, corresponding to an interacting acetylene molecule.

5.3 Results

5.3.1 Percent on surface

A key feature of soot surface analysis involves understanding the proportion of potential reactive sites in the particle that are accessible to a species interacting with the surface. Figure 5.2 shows the percentage of each atom or site type present on the particle surface as a function of the particle diameter. This allows the prediction of the number of reactive surface sites based on particle size, given the total number of atoms or sites in the system. These results are shown for solid-like configurations, and similar trends are seen for higher temperatures, with percent values increasing by $3-4\%$ from solid-like to liquid-like morphologies.

All particles follow a size-dependent trend for surface C atoms and ZZ sites, decreasing from around 55 to 12% and 35 to 5%, respectively. These curves follow a reciprocal diameter trend, indicating that the surface area to volume ratio captured in these values is a controlling factor. This shows that these values are independent of molecular composition, and thus 'real' soot particles can be simplified as homogeneous models. The percent values of H atoms on the surface are significantly higher than C atoms since the solid-like configuration of PAH clusters promotes π-stacking where the molecule edges face outwards, seen for a small heterogeneous particle in Figure 5.3. The proportion of H atoms on the surface shows an added dependence on molecule type, with larger constituent molecules providing larger percentages of surface H values, indicated by an arrow in Figure 5.2(b). This spread can be attributed to differences in stack sizes for the different molecule types. The dissociation energy required to split a PAH stack into two stacks increases with the increasing molecule size, so that stacks are favoured until a length approximately equal to the constituent molecule length [267]. Therefore a particle containing smaller molecules contains a larger number of short stacks, corresponding to increased disorder and more hydrogen atoms hidden in the bulk, than a particle of the same diameter containing larger molecules. The percentage of FE sites on the particle surface is influenced by the presence of different molecule sizes, with

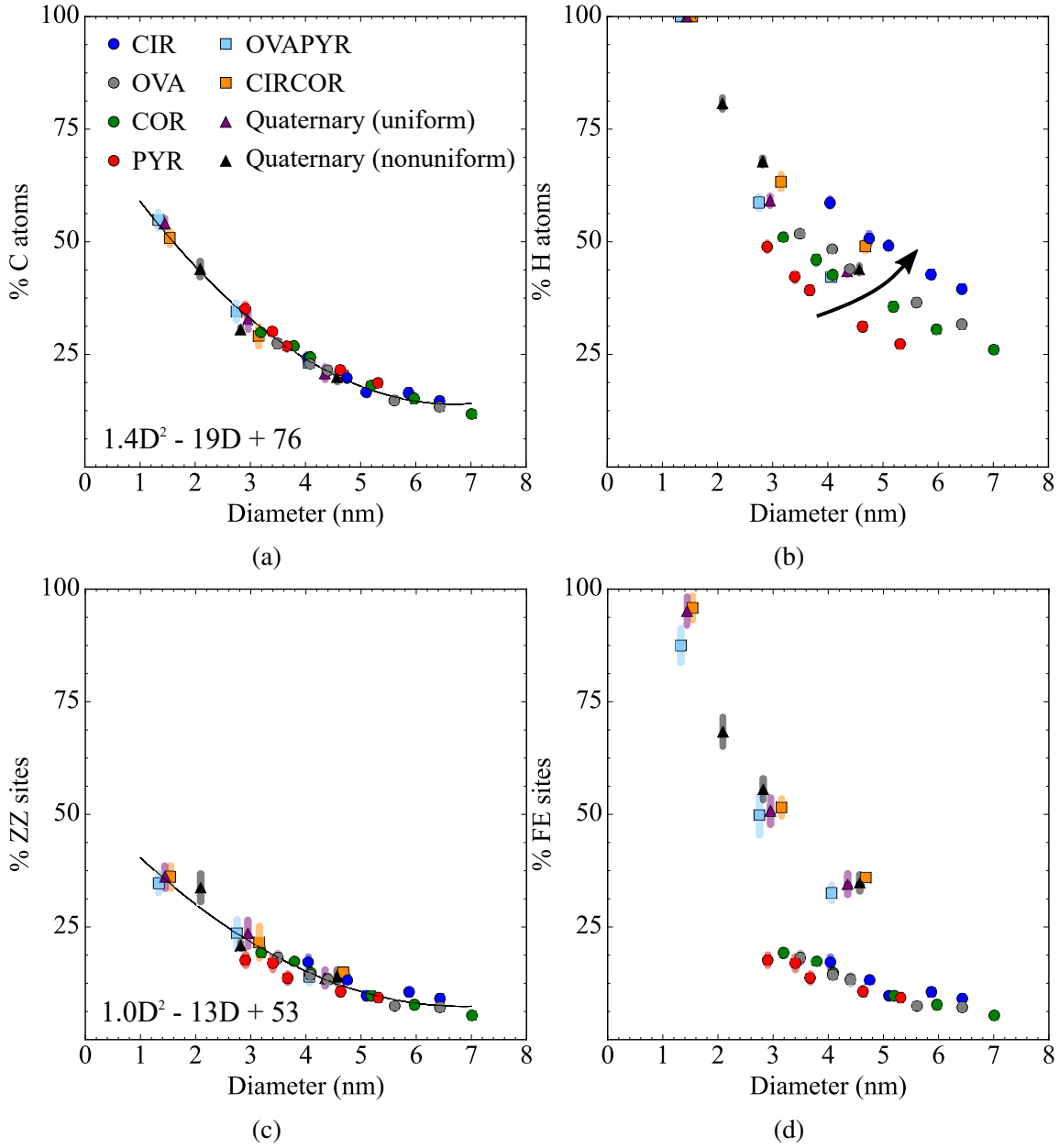
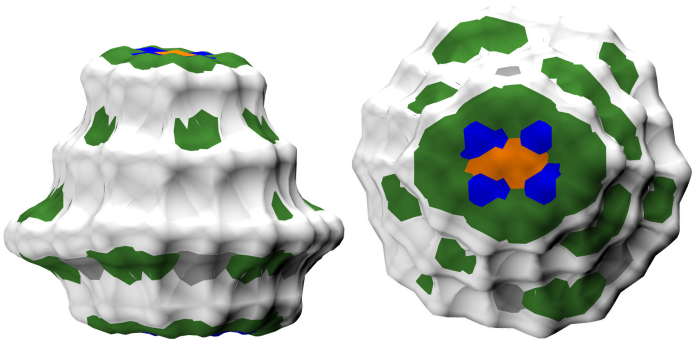


Fig. 5.2 Percentage of (a) carbon atoms, (b) hydrogen atoms, (c) ZZ sites, and (d) FE sites present on the particle surface for homogeneous and heterogeneous particles at solid-like configurations. Shaded areas indicate the standard deviation. The arrow in (b) indicates the effect of increasing molecule size. Homogeneous particle values are taken from Chen et al. [60].

heterogeneous particles presenting significantly higher values ($33-96\%$ for the particle sizes studied here) than homogeneous systems ($5-19\%$). This can be largely attributed to the molecule stacking behaviour. In homogeneous particles the PAHs form aligned stacks and



(a) Side view (b) Top view

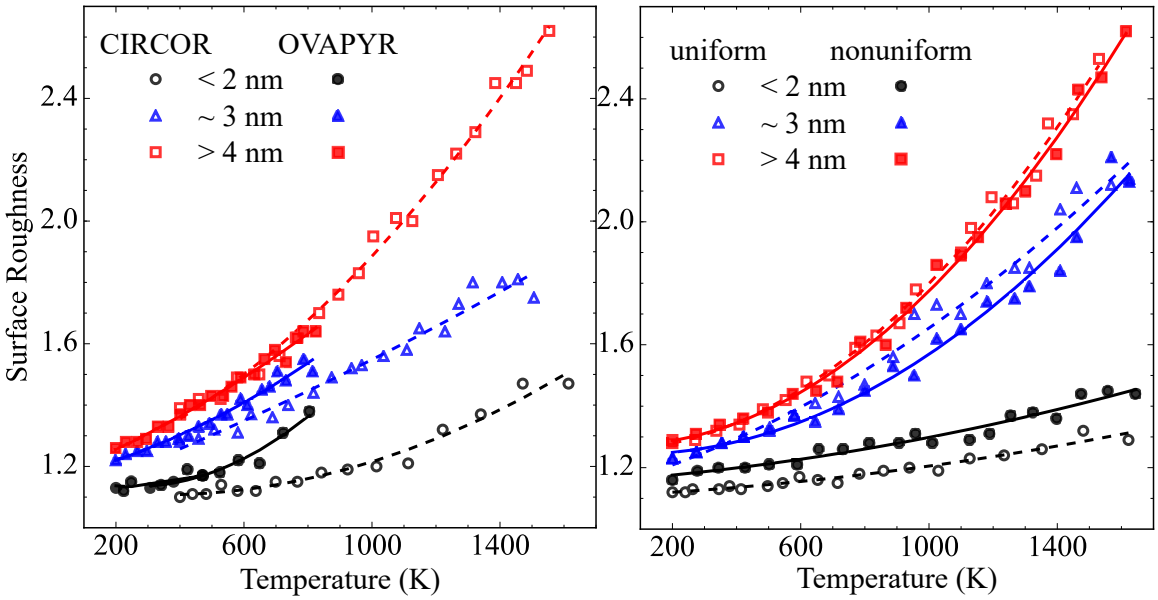
Fig. 5.3 Snapshots of the probe-accessible surface of a 1.45 nm quaternary (uniform) particle in its solid-like state. FE sites, internal ZZ carbon atoms, hydrogen atoms, and IC atoms are coloured green, blue, white, and orange, respectively. This particle has 100% of the H atoms, 54% of the carbon atoms (72% of the edge, 19% of the non-edge), 95% of the FE sites, and 36% of the ZZ sites on the particle surface.

therefore the carbon contributions come predominantly from the molecules capping the ends of the stacks. Heterogeneous particles also form π stacks but the different molecule sizes allow greater accessibility of FE sites for heterogeneous particles compared to homogeneous stacking. For example, the top view of the small heterogeneous particle in Figure 5.3 shows how FE sites on the larger molecules contribute to the particle surface so that 95% of the FE sites are found on the surface. In contrast, the ZZ sites on the surface show the same trends for both homogeneous and heterogeneous particles, again with a typical surface area/volume curve. This is because the molecule size differences are not large enough to reveal the central carbon atoms required for the 3-atom ZZ sites. For example, Figure 5.3 shows that the only inner EC atoms (shown in blue) accessible to the surface probe for the smallest uniform quaternary particle are within the top and bottom molecules in the stack. It is also worth noting that since each molecule contains six FE sites, the percentage of FE sites is dependent on the total number of molecules rather than the particle diameter, seen by the grouping of heterogeneous particles FE site percentages across diameters.

5.3.2 Particle roughness

Particle roughness is calculated as the ratio of the SES surface area to the spherical-equivalent surface area, shown for all heterogeneous particles in Figure 5.4. In all cases roughness increases consistently with temperature. This is unsurprising because increasing thermal energy causes molecules to be more mobile and change from their low temperature config-

uration of parallel stacks to a more disordered configuration in which surface pockets are present [60].



(a) Binary particles (b) Quaternary particles

Fig. 5.4 Surface roughness of binary and quaternary particles with increasing temperature. The smallest particles (diameter < 2 nm) are shown with black circles, intermediate particles (diameter $\approx$ 3 nm) are blue triangles, and the largest particles (diameter > 4 nm) are red squares. Filled and open symbols refer to OVAPYR and CIRCOR particles, respectively, for the binary cases and uniform and nonuniform molecule ratios, respectively, for the quaternary particles. Lines are provided to guide the eye.

Although no single relationship can describe the surface roughness of all particles across temperatures, it is clear that the particle size is an important factor: as the particle diameter increases, its roughness increases. In other words, the larger the particle, the less spherical it is. The proportion of molecules is less important, and the roughness of binary or quaternary particles (either in uniform or nonuniform molecular ratios) is similar. It is interesting to note that at smaller particle sizes, the constituent molecule type provides some contribution as well. This is seen in the increased roughness of the ≤ 3 nm OVAPYR particles compared to their CIRCOR counterparts despite their smaller particle size. This is likely due to the higher rim-to-plane area and elliptical shape of OVA and PYR molecules. All heterogeneous particles follow the same trends and the molecular composition does not play a significant role. This suggests that at low temperatures 'real' soot particles would have a surface area that is 10-30% higher than a spherical approximation would provide, and may be closer to 160% higher for particles > 4 nm at high temperatures.

5.3.3 Surface composition

Temperature has a direct effect on soot particle morphology and therefore influences surface reactivity. Figure 5.5 shows χ_i values (calculated as the number of i reactive sites on the surface over the total SES area) for all particle types and sizes studied here. Scatter points show individual heterogeneous particles, with lines indicating the overall trends. The behaviour of homogeneous particles are provided with dashed trend lines. All systems show the same surface density trends with increasing temperature: χ_C and χ_{ZZ} increase, while χ_H and χ_{FE} decrease. These temperature-dependent behaviours can be explained by the change from solid-like to liquid-like particle morphology and the consequential contribution of different molecule types.

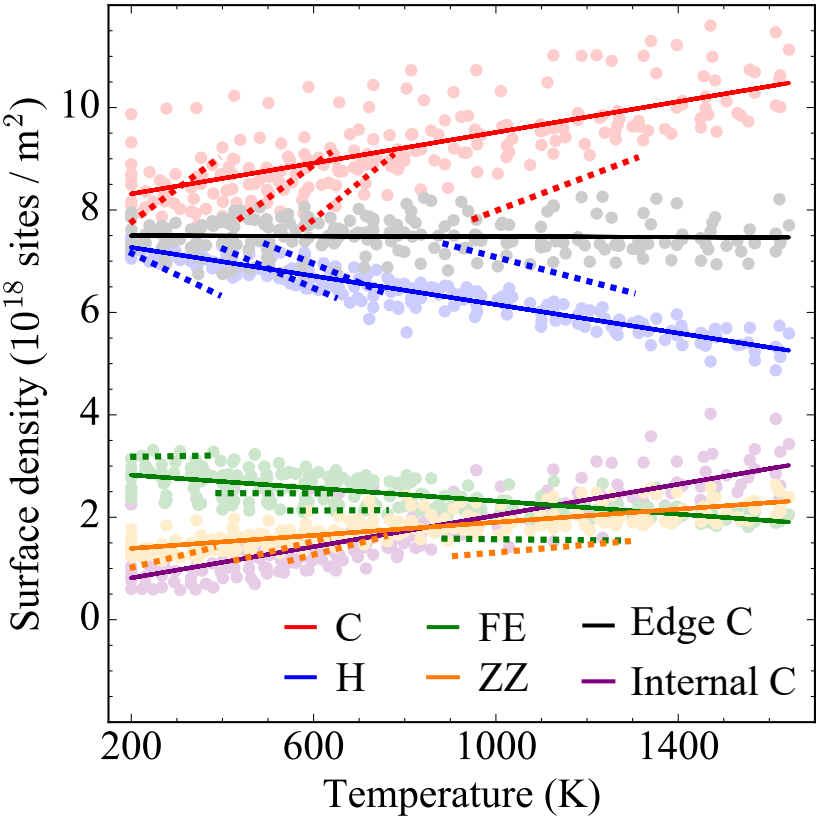


Fig. 5.5 Surface densities of atom types and reactive sites across a range of temperatures for all heterogeneous particles. Surface densities of homogeneous particles, taken from Chen et al. [60], are shown with dashed lines. All lines are provided to guide the eye.

The decrease of χ_H with increased temperature, and the corresponding increase in χ_C, are expected consequences of the morphology changes. Low temperature particles consist of stacked PAHs with molecule edges aligned with the particle surface, so that hydrogen atoms

make up a significant portion of the surface. At increased temperatures, the stacks break up and the molecule planar surfaces are exposed to the surface, providing a significant increase in particle surface area of which the majority is contributed from IC atoms. This is visible in the snapshots of solid-like and liquid-like particles in Figure 5.1. The χ_C and χ_H are primarily dependent on temperature, with consistent values across particle diameters, molecule sizes, and component ratios. For homogeneous particles, separate trend lines are present for each of the molecule types as a consequence of their temperature-dependent behaviour. These homogeneous particle χ_i value differences are a result of reporting as a function of the absolute temperature, since molecule mobility at a given temperature is dependent on the molecule size. Describing these properties with reduced temperature values would produce aligned χ_i values, but this is unfeasible at the moment since heterogeneous melting points for comparison are not known. Mixing molecule sizes in the heterogeneous particles provides single moderate trends.

χ_C shows a relatively large variance and so the contributions from carbon types are investigated further. Isolating the surface density contributions shows that the increase in χ_C with temperature is driven by the IC atoms, while χ_{EC} does not change significantly with temperature. To be clear, the number of sites on the surface increases for all site types as temperature increases (except when the percent of surface H atoms is already at 100% for some of the < 2 nm particles). However, their magnitude increase compared to the increase in particle surface area with temperature determines the density temperature response. In this way, χ_{EC} remains consistent with temperature since the increase in EC atoms corresponds to the increase in surface area.

Interestingly, χ_{EC} is not affected by temperature or particle size, and can therefore be described by the molecular composition alone. As seen in Figure 5.6(a), the contribution of EC atoms to the surface decreases as the constituent molecule size increases, with homogeneous circumcoronene, ovalene, coronene, and pyrene particles presenting values of approximately 3.5×10^{18}, 4.6×10^{18}, 5.2×10^{18}, and 6.7×10^{18}, respectively. A weighted average of these values provides a good prediction of heterogeneous particles examined in this chapter (for example, a CIRCOR particle is an average of CIR and COR values), leading to the simple linear trend $\chi_{EC} = -4.9L + 10$ shown in Figure 5.6(b). This relationship allows the estimation of the density of edge carbons on the particle surface from an average molecule length value, L, such as that determined by fringe analysis of high resolution transmission electron microscopy images. As an example, Alfè et al. [10] gives an average fringe length value of 0.80 nm for young soot produced in a benzene flame. From this, it can be estimated that these particles have a surface density of 6.1×10^{18} atoms of edge carbon per m^2.

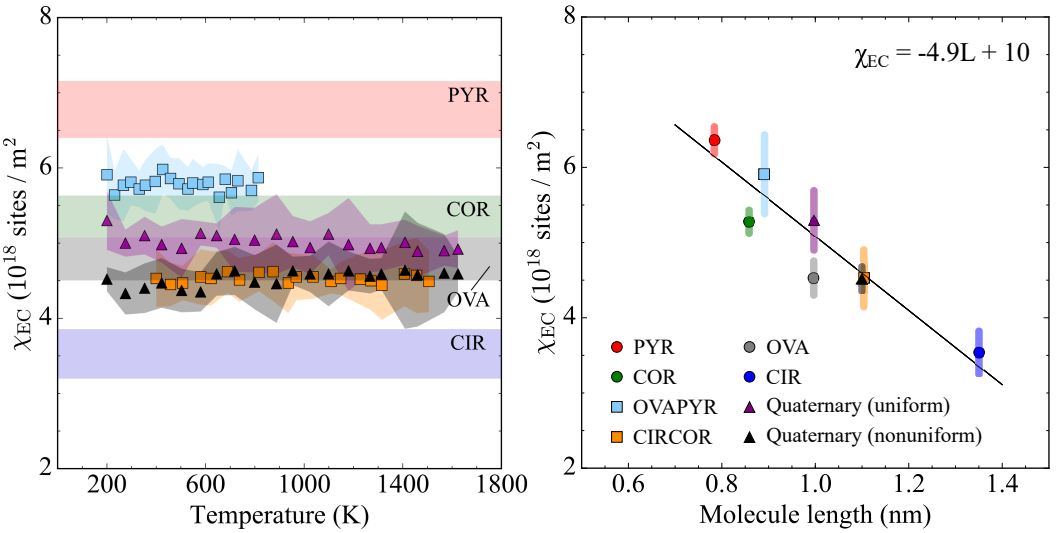


Fig. 5.6 Edge carbon atom density on the particle surface, χ_{EC}, as a function of the temperature (a) and average molecule length, L (b) for particles with a diameter of 3 nm.

The different temperature-independent densities corresponding to the FE sites for the four homogeneous particle compositions in Figure 5.5 suggests that χ_{FE} would show similar behaviour. This is indeed true, as seen in Figure 5.7, although there is a slight decrease in χ_{FE} with increasing temperature due to the increasing contribution of large molecules to the surface. The resulting linear trend $\chi_{EC} = -2.5L + 5$ relates the surface density of FE sites with the average molecule length.

5.3.4 Molecule type contributions

Taking a closer look at the individual molecule types, the same increasing and decreasing atom and site trends seen in Figure 5.5 are observed with increasing molecule size. This suggests that there is an increasing surface contribution from larger molecules on the surface as temperature increases. Figure 5.8 shows the contributions of different molecule types to the total carbon and hydrogen atoms present on the surface for particles with a diameter of approximately 3 nm, at low and high temperatures corresponding to solid-like and liquid-like particles, respectively. These percent values are normalised by the expected contributions taking into account the proportion provided by each molecule.

At solid-like configurations, the surface carbon atoms experience higher contributions from the small molecules than expected from the molecular proportions. This can be explained by the core-shell configuration at low temperatures seen previously. The two types of bi-

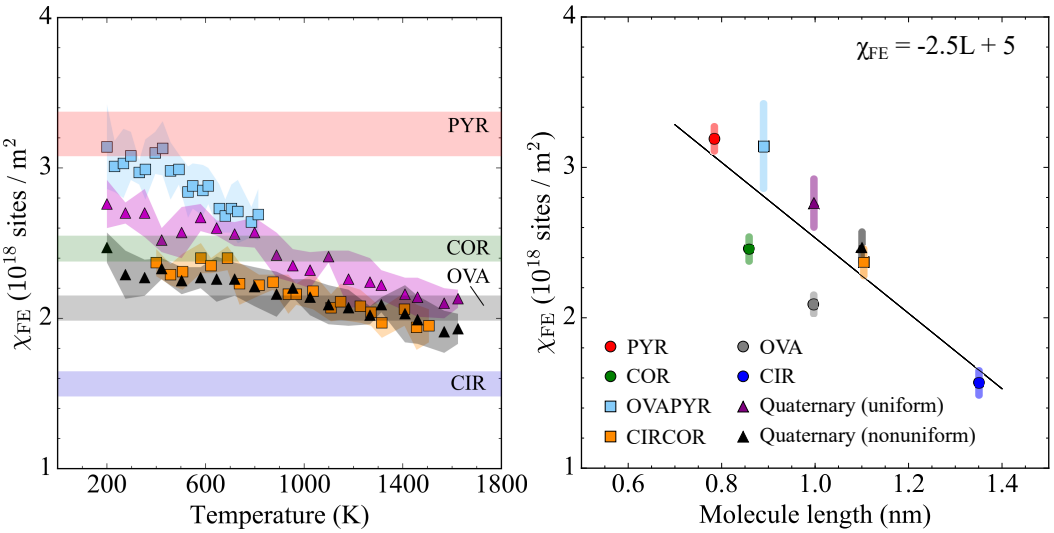
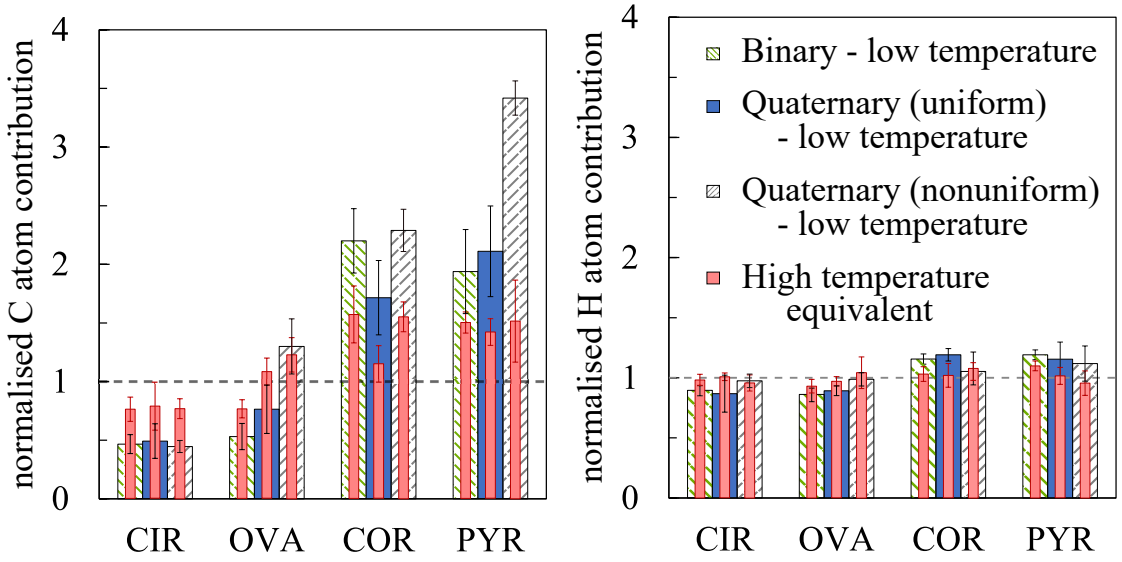


Fig. 5.7 Free-edge site density on the particle surface, χ_{FE}, as a function of the temperature (a) and average molecule length, L (b) for particles with a diameter of 3 nm.

nary particles (CIRCOR and OVAPYR) show half of the expected contribution from the larger molecules (CIR or OVA) and double the expected contribution from the smaller molecules (COR or PYR). The quaternary particles show the same behaviour, with increasing contributions from decreasing molecule sizes. The difference between uniform and nonuniform quaternary particles highlights the presence of small molecules on the particle surface. Nonuniform particles have higher normalised values for the small molecules since the lower proportion of small molecules in these systems would be expected to play a smaller role in the surface composition. These results show that small molecules are present on the particle surface regardless of the distribution of molecules sizes in the particle. At high temperatures the molecule size differences decrease towards their expected molecular contributions, corresponding to a mixed arrangement in which all of the molecule types contribute proportionately to the surface properties. In contrast, all molecule types contribute proportionately to the surface H atoms in both solid-like and liquid-like configurations.

5.3.5 Alpha

The parameter α provides a reactivity factor that accounts for temperature-dependent sterics, due to the increasing entropy and molecule mobility with increasing temperature, which causes molecules to break the stacked structure that exposes hydrogen atoms to the surface and introduces more unreactive basal planes, as well as the increasing contribution of larger



(a) (b)

Fig. 5.8 Molecule type contributions to the carbon and hydrogen atoms on the surface of 3 nm particles. Low temperature values provide solid-like particles, and the corresponding liquid-like particles at high temperatures are shown with inset narrow red bars. All values are normalised by the expected values considering molecule type proportions.

molecules. The parameter α is defined as the area of hydrogen atoms over the total SES area. Figure 5.9(a) shows α values for a series of homogeneous and heterogeneous 3 nm particles containing different molecule types. As with the χ_H values, there are distinct downward trends for each homogeneous particle type with temperature. Interestingly, increasing heterogeneity converges the α parameter to a similar temperature dependence, which can be approximated by a simple linear temperature-dependent relationship. The heterogeneous particles show α values consistent with a weighted average of the molecular components, suggesting that the increasing heterogeneity of 'real' soot particles produces a consistent trend. The same behaviour is seen for particles $>$ 4 nm, shown in Figure 5.9(b).

The work presented in this chapter forms a foundation for understanding how heterogeneity in molecule size influences the surface properties of soot particles. The site composition of molecules, size differences between constituent molecules, molecular partitioning within the cluster, and tendency to form molecular stacks all play a role in the corresponding cluster surface properties, and thus it is expected that molecule types that enhance or disrupt these features would see corresponding changes in cluster surface properties. For example, perhaps the presence of crosslinked molecules will cause a larger disparity amongst molecule sizes, decrease the degree of π-π stacking interactions, or decrease overall molecule mo-

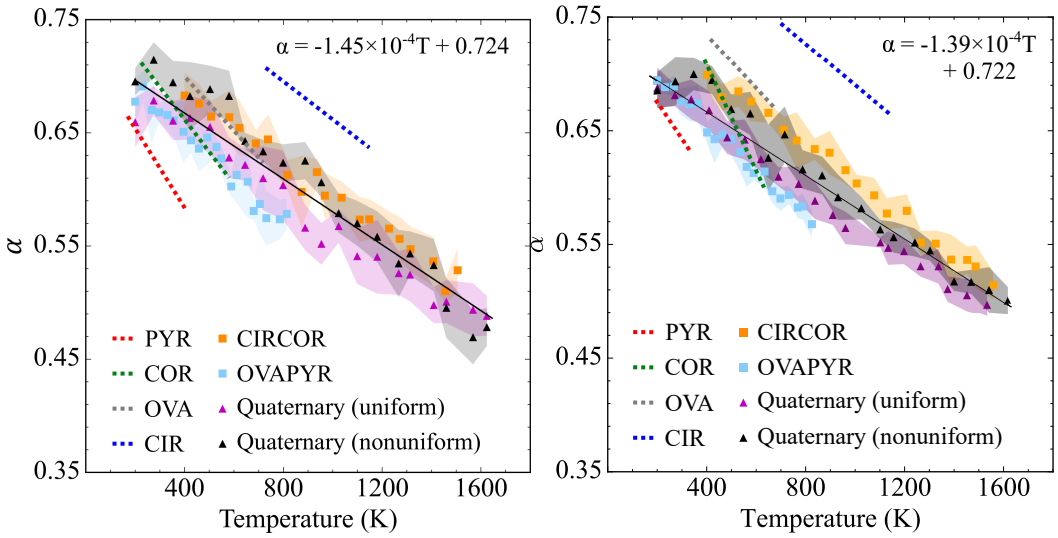


Fig. 5.9 Parameter α for particles with approximate diameters of (a) 3 nm and (b) 5 nm as a function of temperature. Shaded areas show the standard deviation of heterogeneous particle values. Homogeneous particle values, taken from Chen et al. [60], are shown as dashed lines. A single trend line is shown for all of the heterogeneous cases.

bility and degree of core-shell partitioning; these features may be competing factors that influence cluster structure and influence surface properties. Further work should build on these heterogeneous cluster findings to explore the influence of complex molecule types proposed to play a role in soot, such as PAHs with crosslinks, aliphatic chains, asymmetry, or curvature [101, 217], in surface properties. For example, increasing curvature is seen to substantially increase oxidative reactivity but it is unknown to what extent this is due to the individual molecule reactivity or a consequence of an altered nanostructure of particles containing curved aromatic molecules [335].

5.4 Conclusions

PAH particles containing two or four molecule types are studied to provide insight into the effect of molecular heterogeneity on soot particle surface properties. A thorough analysis is conducted across particle sizes (1–5 nm) and temperatures (200–1600 K). Relationships between particle diameter and the percentage of reactive sites on the particle surface are identified, showing that C atom and ZZ site percentages do not change with heterogeneity. Molecule size differences in heterogeneous particles increase the FE sites available for reactions, while H atom percentages increase with constituent molecule size. Particle

roughness increases with temperature and particle size, with particle composition playing a secondary role. The controlling parameters for surface densities of reactive sites are also identified, showing that χ_{H}, χ_{IC}, and χ_{ZZ} are dependent on temperature while χ_{EC} and χ_{FE} can be predicted from the average molecule length. These trends are influenced by enhanced surface contributions from small molecules at low temperatures. Finally, the temperature-dependent behaviour of the parameter α shows a correlation to constituent molecule size, suggesting that the α parameter for heterogeneous particles converges to an average value of the constituent molecules.

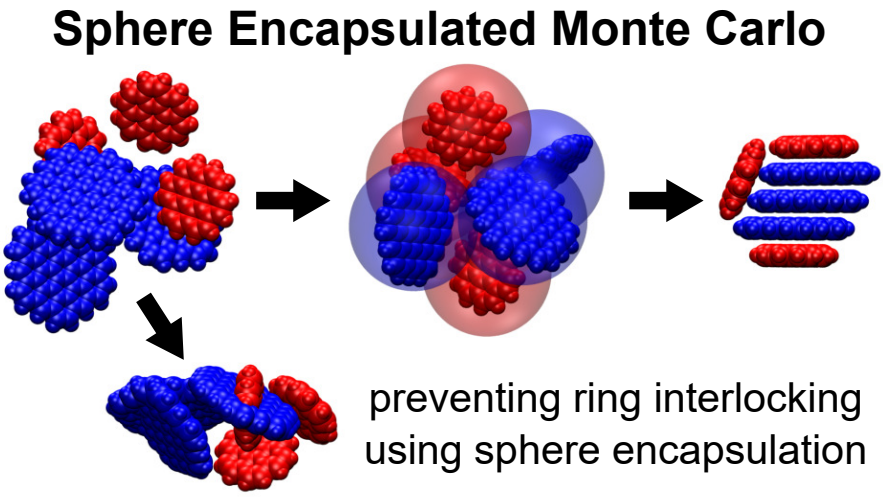
Chapter 6

Optimisation method for large aromatic systems

A simple global optimization approach is introduced that is able to find minimum energy configurations of clusters containing aromatic molecules. The translational and rotational perturbations required in Monte Carlo-based methods often lead to unrealistic configurations within which two or more molecular rings intersect, causing many of the computational steps to be rejected and the optimization process to be inefficient. A modification of the basin-hopping global optimization procedure tailored to tackle problems with intersecting molecular rings is developed. Termed the Sphere Encapsulated Monte Carlo (SEMC) method, this method introduces sphere-based rearrangement and minimization steps at each iteration, and its performance is shown through the exploration of potential energy landscapes of polycyclic aromatic hydrocarbon (PAH) clusters, systems of interest in combustion and astrophysics research. The SEMC method provides clusters that are accurate to 5% mean difference of the minimum energy at a 10-fold speed up compared to previous work using advanced molecular dynamics simulations. Importantly, the SEMC method captures key structural characteristics and molecular size partitioning trends as measured by the molecular radial distances and coordination numbers. The advantages of the SEMC method are further highlighted in its application to previously unstudied heterogeneous PAH clusters.

Highlights:

- Sphere Encapsulated Monte Carlo method is developed to overcome ring interlocking
- Minimum energy configurations are determined at low computational expense
- Applied to clusters beyond the scope of existing methods



The work presented in this chapter was published as a paper in the Journal of Physical Chemistry A by Bowal et al. [42].

6.1 Introduction

Stable arrangements of nanoclusters may provide structural and thermodynamic properties and their determination is an area of great interest in the physical sciences. For example, stable arrangements can help to identify native states of a protein [289], the experimental structure of an atomic cluster [8], or the density of a molecular cluster [323]. The geometric ground state of a system represents the global minimum of its corresponding potential energy surface (PES). Global minimum determination is not a trivial task and strongly depends on the properties of the PES function, which are usually unknown. In addition, for any case of interest there will always be more than one stable configuration, which results in an energy that is a non-convex function of the nuclear coordinates. Scanning across the entire PES to ensure the global minimum is reached is a challenging task at best, and computationally unfeasible for most systems of interest. For example, a cluster containing 13 atoms has at least 1510 minima [91, 273], a cluster containing 75 atoms has 8×10^6 minima [342], and a cluster containing 147 atoms possesses something on the order of $10^{60} - 10^{259}$ minima [328].

Many global optimisation methods have been developed to tackle this problem for a variety of systems. One popular method uses consecutive stochastic displacements and local minimisation steps to efficiently sample a complex PES. This Monte Carlo minimisation method was first used to examine protein folding by Li and Scheraga [203] in 1987 and subsequently generalised and popularised by Wales and Doye [344] as the basin-hopping method. The basin-hopping minimisation method simplifies the search of a complex PES by transforming it to a series of low energy plateaus, called basins [344]. This is done by taking the energy of each sampled configuration as the energy of its nearest local minimum, consequently removing all downhill barriers from the energy landscape while retaining all minima. In order for the searching method to move between basins, all components within the system studied are randomly translated and rotated between each minimisation step. Implementing this large stochastic rearrangement prevents the system from being trapped in a local minima and promotes scanning of the entire PES. The basin-hopping method has been shown to be an effective global optimisation technique that allows for the determination of energy minima through exploration of a small fraction of the existing local minima, and has successfully determined the minimum energy configurations of many atomic clusters [19, 92, 90, 89, 227]. Due to its simple implementation and low computational cost, the basin-hopping method has been applied to many fields [346].

Since its development several decades ago there have been many tools implemented with the basin-hopping algorithm to improve the efficiency of global optimisation for specific

systems, including parallel tempering [243], path memory to avoid revisiting sites [127], order parameters [138, 275], trial operators [274], additional jump moves [169], reseeding, symmetrising, and taboo lists [341]. More details can be found in a recent review on the developments in basin-hopping global optimisation by Wales [343]. The basin-hopping method is well-suited to atomic systems or those in which molecules can be represented as spherical or lumped superatomic components [150, 93, 312, 193] but is less efficient when evaluating complex molecular systems, such as clusters containing aromatic molecules, due to the required stochastic rearrangement steps. This means that the choice of moves to perform the random perturbation steps is critical. For example, molecular clusters and biomolecules have been studied successfully using local rigidification and group translation and rotation moves within the basin-hopping optimisation framework [305, 245].

Clusters of polycyclic aromatic hydrocarbons (PAHs) are systems of intense interest within the fields of combustion [39, 265], aerosols [78], and astrophysics [33]. However they present a challenging system for global optimisation methods due to their strong molecular anisotropy. The large random rotation and translation steps required for stochastic basin-hopping sampling cause frequent intersections and ring interlocking between aromatics, leading to many unrealistic overlapping geometries and thus causing the process to become inefficient [267]. Minimum energy structures of PAH clusters have been explored using basin-hopping methods [323, 267], evolutionary algorithms [22], and other algorithms involving iterative random movements and local minimisations [131, 317], however due to computational inefficiency these studies are limited to small PAH clusters. Recently, a coarse-grained potential was developed that represents each PAH as an ellipsoid [153]. This approach smooths the PES and avoids this sampling problem, although at the expense of interaction accuracy and transferable parameters. This coarse-grained potential was used with the basin-hopping method to evaluate low energy structures of clusters containing up to 20 molecules [153] and provide thermodynamic and aggregation behaviour of clusters containing up to eight PAHs using dynamic modelling methods [152, 151].

PAH clusters containing pyrene, coronene, ovalene, hexabenzocoronene, or circumcoronene have been studied previously using molecular dynamics methods [64, 63, 61, 60, 132]. Previous work exploring heterogeneous PAH clusters described in Chapter 4 used replica exchange molecular dynamics and found that this method was able to provide minimum energy stable configurations for clusters containing up to 100 PAHs, but at a high computational cost due to the large number of parallel simulations required. Interestingly, a partitioning of PAHs was observed within these heterogeneous clusters, with larger PAHs residing near the cluster centre and smaller PAHs remaining around the cluster surface. These REMD simulations required a radial position potential to pull evaporated PAHs back to the cluster.

Although this was shown to play a role in the high temperature replicas without significantly impacting the low temperature replicas of interest, it is valuable to verify the molecule-size partitioning using an independent Monte Carlo approach which requires no such position potential. In addition, an improved sampling method would allow for the arrangements of aromatic clusters to be robustly sampled with significantly less computational expense than REMD, allowing for larger and more complex systems to be explored.

This chapter presents a novel method of implementing the basin-hopping minimisation method for clusters containing molecules with highly anisotropic geometries, the Sphere Encapsulated Monte Carlo (SEMC) method. Using the basin-hopping Monte Carlo framework, the SEMC method incorporates additional sphere-based rearrangement and minimisation steps to provide efficient sampling of complex systems that have previously presented challenges. The usefulness of this method is displayed through its application to clusters composed of polycyclic aromatic hydrocarbons. The low energy configurations of homogeneous and heterogeneous PAH clusters are identified and compared with recent work using advanced and costly molecular dynamics simulations.

6.2 Methods

The SEMC method provides a computationally efficient means to explore the complex PES of PAH clusters in order to evaluate their stable molecular arrangements. Following the basin-hopping framework, the SEMC method consists of a series of random perturbations and minimisations in order to locate minimum energy structures. The key novelty of the SEMC method lies in the use of spheres to represent PAH molecules in order to improve the efficiency of the rearrangement steps.

The procedure of the SEMC method is as follows, also shown as a flow chart in Figure 6.1. An initial cluster with an energy of E_{initial} is generated (more details provided below). Each PAH is then represented by a sphere with a diameter corresponding to the diameter of the largest PAH in the cluster (the importance of having spheres of equal size is discussed later).

The position of each sphere is randomly translated by a three-dimensional vector generated from a Gaussian distribution with a mean of zero and a standard deviation of 2.5 Å. This rearranged system is minimised using a computationally efficient Lennard-Jones (LJ) potential to provide a minimum energy configuration of spheres. Following the description of the LJ potential provided in Equation 3.2, σ is the distance in Åbetween the molecule centre to the most distant hydrogen centre and the van der Waals radius of hydrogen and ε is an arbitrary

value in eV that provides a sufficiently deep potential energy well. This potential does not provide the detailed PAH interactions of the isoPAHAP potential but allows for a computationally efficient determination of the minimum energy configuration of a cluster containing spheres, suitable for the rearrangement steps. Minimisation is conducted using the conjugate gradient method for a maximum of 10^4 iterations or until the convergence criterion is fulfilled. This step adjusts the randomly positioned spheres so that they form a compact cluster in which the spheres do not overlap and plays a role in reducing the configurational space being searched by this Monte Carlo global optimisation algorithm. The use of spheres for the encapsulating shape was selected in order to eliminate intersection or other unfavourable interactions between component molecules during the translation and rotation steps. The SEMC method can be easily generalised to use non-spherical encapsulation shapes, such as ellipsoids or cubes or more complex volumes. However a random rotation of any molecule presents a spherical volume projection if rotation is allowed in all dimensions and thus the use of spheres allows for the most efficient use of space during the encapsulation steps.

The PAH molecules are then reinserted into their corresponding spheres and are randomly rotated using Euler angles that produce uniformly distributed orientations [191] with a standard deviation of 1 radian. Due to the use of spheres previously, the PAHs are allowed to freely rotate without intersecting with neighbouring molecules. The molecular cluster is then minimised using the isoPAHAP potential to generate a minimum energy molecular structure. The isoPAHAP potential is a PAH-specific potential developed from high accuracy symmetry-adapted perturbation theory calculations [324] and has been used successfully in a number of previous studies on PAH systems [131, 326, 63]. This molecular energy minimisation uses the limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm [48] for a maximum 10^5 iterations or until the maximum force acting on any atom in the system converges to 0.01 eV/Å.

As in the basin-hopping method, the energy of this arrangement is compared with the most recent cluster minimum energy and moves are accepted or rejected based on the Metropolis criterion [225, 144]. Hence if the new cluster energy, E_{new}, is lower than the previous cluster minimum energy, E_{current}, then the corresponding new configuration is accepted as the updated current reference configuration and a new iteration begins from this new system. Otherwise, the new configuration is accepted with the probability $\exp[(E_{\text{current}} - E_{\text{new}})/(k_B T)]$. Here, k_B refers to the Boltzmann constant and T to the simulation temperature. A temperature of 10 K is used for the smallest clusters made up of fewer than 5 molecules, 500 K for the larger clusters containing up to 32 molecules, and 1000 K for the largest clusters containing 100+ molecules. The selection of these temperatures will be discussed further.

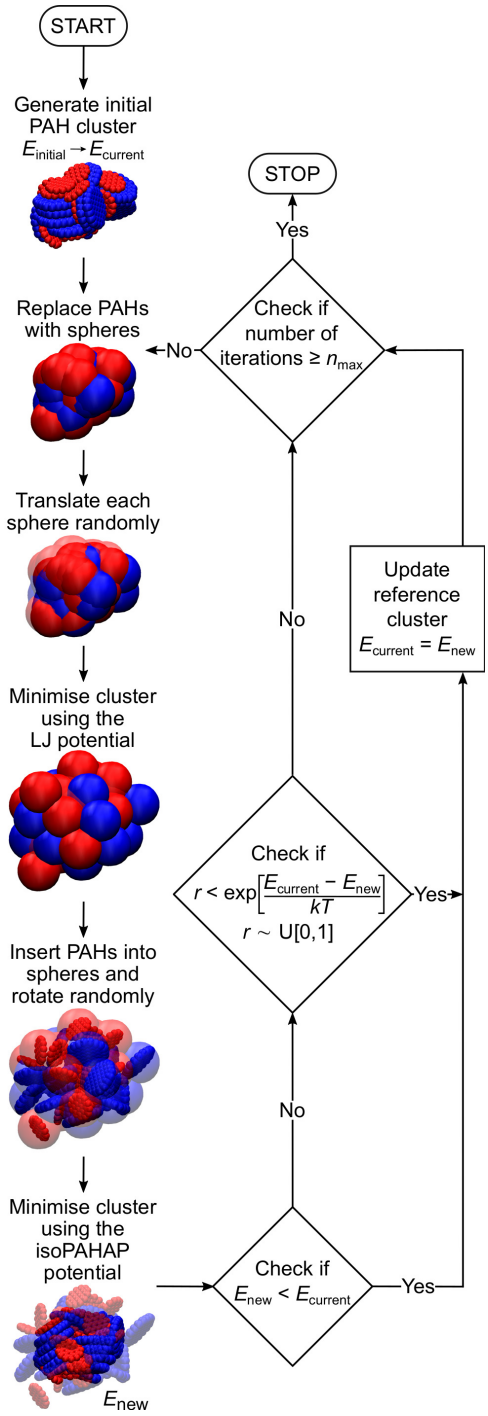


Fig. 6.1 Flow chart of the SEMC method. The cluster images show a $CIR_{16}COR_{16}$ cluster, where CIR molecules are shown in blue and COR molecules are shown in red.

Note that the initial cluster of PAH molecules is generated using a similar process as above: Spheres are randomly distributed within a large spherical volume with a radius of 25 Å. This

cluster undergoes a LJ minimisation, the PAHs are inserted into their corresponding spheres and then rotated, and the molecular cluster is minimised using the isoPAHAP potential to produce an initial cluster with an energy of E_{initial}.

To implement the SEMC method, purpose-made Python scripts were written to perform sphere encapsulation and Monte Carlo steps, which call upon the software package EON [254, 69] to execute sphere- and molecule-based energy minimisations. EON is a freely available atomic simulation software that was modified in this case to use the LJ and isoPAHAP intermolecular potentials.

Heterogeneous PAH clusters containing different combinations of pyrene (PYR), coronene (COR), ovalene (OVA), and circumcoronene (CIR) are simulated. These molecules are selected because they span the range of PAH sizes relevant to soot particles [39] and interstellar media [322]. The notation $\text{AAA}_\text{x}\text{BBB}_\text{y}$ will be used to describe the clusters studied, where AAA and BBB refer to the molecule types and x and y give the number of molecules. All clusters undergo at least three independent repetitions each with $n_{\max} = 1000$ iterations.

Average radial distances and coordination numbers are calculated to provide insight into molecular arrangements and cluster structure. The average radial distance of molecule type k, r_k, is calculated using the following equation:

$$r_k = \frac{1}{N_k} \sum_{m \in I_k} R_m \tag{6.1}$$

where N_k is the number of molecules within $k \in \{\text{small}, \text{large}\}$, referring to the small and large molecule types, I_k is the set of molecule indices within k, and R_m is the radial distance between the cluster geometric centre and the geometric centre of molecule m.

The coordination number of molecule type k, CN_k, provides the average number of molecules within a distance $R^{(\text{cutoff})}$. This is calculated as the average of the cumulative radial distribution function integrated over the centre of the cluster (set as the origin at 0) to a maximum cut-off radial distance, $R^{(\text{cutoff})}$:

$$CN_k(R^{(\text{cutoff})}) = \frac{1}{N_k} \sum_{m \in I_k} \int_0^{R^{(\text{cutoff})}} \sum_{j=1}^{N_{\text{total}}} \delta(D_{mj} - R) dR \tag{6.2}$$

$k \in \{\text{small}, \text{large}, \text{total}\}$ refers to the small molecules, large molecules, and all molecules within the cluster, respectively. D_{mj} is the distance between the geometric centres of molecules m and j. $R^{(\text{cutoff})}$ corresponds to the radius of the molecule type of interest, so $R^{(\text{cutoff})}_{\text{PYR}} = 0.46$ nm, $R^{(\text{cutoff})}_{\text{COR}} = 0.47$ nm, $R^{(\text{cutoff})}_{\text{OVA}} = 0.59$ nm, and $R^{(\text{cutoff})}_{\text{CIR}} = 0.71$ nm.

CN_{total} values use the radius of the larger molecule in the system considered. These provide *CN* values from zero to two, corresponding to an isolated molecule and a molecule sandwiched between two others, respectively.

In this chapter, homogeneous clusters containing $2-32$ coronene molecules are simulated to provide an assessment of the SEMC method in a system previously studied using several optimisation methods, including basin-hopping minimisation [267, 153] and an evolutionary algorithm [22]. Following this, investigations into the low energy cluster structures of heterogeneous clusters containing up to 100 molecules are performed. Relevant results are compared to recent simulations of PAH clusters using REMD (described in Chapter 4) and the mean absolute differences between the methods are reported, illustrating the accuracy and efficiency of the SEMC method.

6.3 Results and discussion

Key components of the SEMC method development, including evaluation of fundamental interactions with benchmark tests of COR clusters, selection of the temperature parameter, and important considerations for heterogeneous cluster systems are discussed. This is followed by qualitative and quantitative results for heterogeneous PAH cluster systems, presented first in detail for a representative cluster and followed by a summary of results from all clusters. A direct comparison is made with previous molecular dynamics calculations. Finally, the results from several other systems are given, illustrating the capability of the SEMC method to predict the properties of larger and more complex clusters than was previously feasible using other computational methods.

6.3.1 SEMC method development

Coronene cluster benchmark

Coronene is a seven-ringed PAH often studied as a model representation of graphene, since it is the smallest PAH that presents a central carbon ring chemically bound only to other carbons [257, 278, 373]. COR clusters also provide benchmark systems for studying PAH interactions and are useful for testing the reliability of novel optimisation methods. The SEMC method is applied to homogeneous COR clusters containing 2–32 molecules. This allows the minimum structures produced using this novel method to be compared with those determined by other established global minimisation methods: the basin-hopping

method [344] using all-atom [267] or coarse-grained potentials [153] and an unbiased hybrid evolutionary algorithm method [209] using an all-atom potential [22], as shown in Figure 6.2. All four methods show similar energy trends as a function of the size of the COR cluster and present similar minimum energy structures. However, since each method uses a different intermolecular potential to describe the atomic interactions, detailed comparisons of each cluster energy cannot simply be ascribed to the global optimisation method used. In particular, the lower energies presented by the two studies using the basin-hopping method [267, 153] are not due to the optimisation method but are a consequence of the use of generic LJ parameters, which are known to overestimate the stacking energy of PAHs and influence cluster morphology [252]. This has been corrected by potentials such as the isoPAHAP potential [324] and the improved LJ potential [22] which are parametrised for PAH interactions. The energy shift caused by an accurate PAH potential is highlighted with a purple arrow in Figure 6.2.

A more detailed evaluation of the minimum energy values determined by the evolutionary algorithm with an improved LJ potential compared to calculations using the SEMC method with the isoPAHAP potential also highlights the strength of the latter in capturing all PAH interactions. Good agreement between the two methods for small clusters (dimer, trimer, etc.) suggests that sandwich-like interactions in which the molecules interact through π - π bonding are well captured by both methods. In both cases, for a cluster containing six or more molecules, the structural motif changes from a single stack to several smaller stacks, as explained in detail by Bartolomei et al. [22]. In these larger cluster configurations T-shaped interactions play a more significant role and thus the slight energy differences between the evolutionary algorithm with an improved LJ potential and the SEMC method with the isoPAHAP potential are likely due to the description of CH - π interactions. The lower energies found by the SEMC method are unsurprising since the development of the isoPAHAP potential included careful parametrisation of T-shaped interactions to high accuracy symmetry-adapted perturbation theory based on density functional theory calculations [324]. For example, the minimum energy cluster containing ten COR molecules presents the same structure using both the evolutionary algorithm with an improved LJ potential and the SEMC method with the isoPAHAP potential: two equal-sized curved stacks in a handshake arrangement, shown in Figure 2. T-shaped interactions are significant in this structure and differences in how they are described by the two intermolecular potentials is the cause of the energy difference between the two methods. Thus the SEMC method is able to determine COR cluster structures in good agreement with literature results, and the use of the isoPAHAP potential provides accurate energies for clusters containing T-shaped interactions.

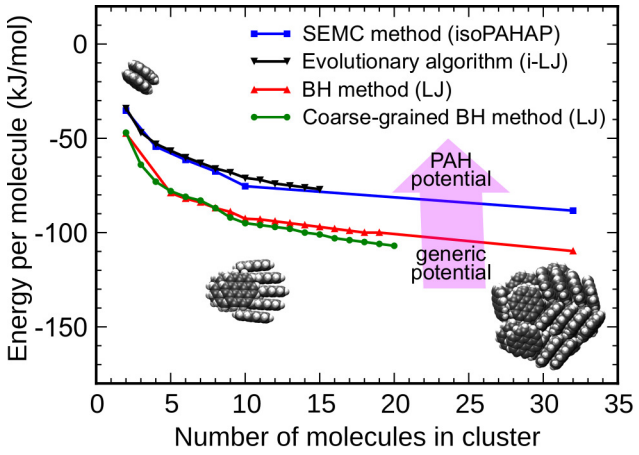


Fig. 6.2 Energies of homogeneous PAH clusters containing 2–32 coronene molecules obtained with the SEMC method using the isoPAHAP potential and published global optimisation methods. The values obtained by the evolutionary algorithm method using an improved Lennard-Jones potential come from Bartolomei et al. [22], the basin-hopping method using a Lennard-Jones potential come from Rapacioli et al. [267], and the coarse-grained basin-hopping method using a Lennard-Jones potential from Hernández-Rojas et al. [153]. The purple arrow highlights the energy shift caused by using PAH-specific interaction parameters compared to generic LJ parameters which are known to cause enhanced binding. Cluster snapshots shown are from the SEMC method.

Temperature parameter selection

Temperature is a key free parameter in the SEMC method, controlling the degree to which higher energy configurations are accepted and thus determining the movement between minimum energy basins as the simulation progresses. For relatively small values of the temperature parameter, the system may exhibit the tendency to remain stuck in a current local minimum. On the other hand, too high values of the temperature parameter are known to promote many higher energy configurations to be accepted, thus promoting significant movement across energy basins without searching within minimum energy wells. Selection of the temperature parameter is dependent on system size, and therefore an assessment is conducted to evaluate appropriate values for the clusters of interest. The number of steps required to obtain the minimum energy for different cluster sizes is assessed for at least three independent simulations starting from different initial configurations and the minimum energy results for each case are shown in Table 6.1. It is seen that for a small homogeneous cluster, a low temperature of 10 K is favourable, while the higher temperatures of 500 K and 1000 K are able to locate the minimum energy values more quickly in the 32 molecule and 100 molecule clusters, respectively. Therefore a temperature of 1000 K is used for the

largest clusters, 10 K is used for the smallest clusters, and 500 K is used for the intermediate clusters containing more than 5 but fewer than 100 molecules. These selected temperatures provide satisfactory movement across the PES and convergence toward a minimum energy configuration for the clusters investigated. In most cases, the minimum energy is determined within 1000 iterations at all temperatures considered. While a more rigorous temperature selection process is possible and may result in a more efficient programme for specific systems of interest, this is out of the scope of this thesis.

Table 6.1 Number of iterations of the SEMC method required to obtain minimum energies (kJ/mol) using different values of the temperature parameter (K). Energies and iterations are provided as ranges, taken from at least three independent simulations conducted for each case. The average iteration value is provided in parentheses.

Cluster	Temperature	Minimum energy	Iterations
COR_4	10	-217 to -218	317 to 959 (580)
	500	-201 to -202	969 to 999 (989)
$CIR_{16}COR_{16}$	10	-4630 to -4660	154 to 293 (224)
	500	-4640 to -4710	277 to 866 (573)
$OVA_{50}PYR_{50}$	500	-9890 to -9910	530 to 769 (631)
	1000	-9860 to -9910	6 to 447 (197)

Process development for heterogeneous clusters

Initial SEMC simulations used spheres with diameters corresponding to the size of the molecule they were replacing. For heterogeneous PAH clusters containing more than one type of PAH, this caused the LJ minimisation step to act on a cluster containing spheres of two different sizes. It was observed that these heterogeneous LJ clusters favoured an arrangement in which the smaller spheres were located in the cluster centre in order to equalise surface and bulk bonds. The favouring of a core-shell configuration was reported in other heterogeneous cluster studies containing spherical components of different sizes [236, 90]. In the SEMC process, a size-dependent partitioning of spheres is undesired since the purpose of the sphere encapsulated steps is to randomly rearrange the PAH cluster. Thus the SEMC method uses equally sized spheres, corresponding to the size of the largest PAH in the cluster, eliminating any partitioning behaviour within the rearrangement steps. This modification promotes lower energy cluster configurations, as shown in Figure 6.3 for a $CIR_{16}COR_{16}$ cluster. This indicates that a cluster containing smaller PAHs in the core is less energetically favoured, in contrast to the results seen for binary LJ sphere systems.

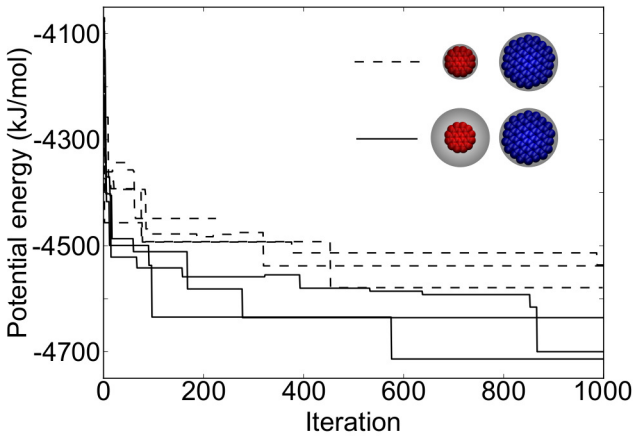


Fig. 6.3 Accepted energies as a function of iteration for the $CIR_{16}COR_{16}$ cluster. Three cases initialised with different configurations are given for both simulations using spheres of different sizes in the rearrangement step (shown with dashed lines) and simulations using a uniform sphere size (shown by solid lines).

6.3.2 Heterogeneous PAH clusters

The SEMC method is used to determine the low energy structures of a number of heterogeneous PAH clusters: CIR_8COR_8, OVA_8PYR_8, $CIR_{16}COR_{16}$, $OVA_{16}PYR_{16}$, $CIR_{50}COR_{50}$, and $OVA_{50}PYR_{50}$.

The minimum energy arrangements of these molecular clusters is of interest, in particular to evaluate if the SEMC method is able to capture the interactions between planar aromatics and the effect of two molecule types within a single cluster. To do this, the cluster configurations were assessed qualitatively and using several quantitative measures. To aid in understanding this analysis, the results are first reported and discussed in detail for an intermediate-sized representative PAH cluster, CIR_8COR_8, followed by results from all clusters studied.

Detailed analysis of representative heterogeneous cluster

Figure 6.4 shows the accepted cluster energies over the course of an SEMC simulation evaluating a CIR_8COR_8 cluster, as well as images of the cluster configurations at low energy points. In this case, the accepted cluster energies decrease steadily as the simulation progresses and the lowest energy configuration is found near to the end, providing a clear example of progressive cluster arrangements at decreasing energies. At iteration 1, the cluster has a potential energy of -1841 kJ/mol and the cluster shows some stacking of the molecules although it is not compact and the two molecule types are mixed. As the simulation progresses through iterations 49 (-2137 kJ/mol) and 575 (-2142 kJ/mol), these lower energy

clusters show increasingly well-stacked structures with molecules interacting more strongly with other molecules of the same type such that the two molecule types separate. The final lowest energy cluster at iteration 926 (-2189 kJ/mol) contains a single central stack of CIR molecules surrounded by smaller stacks of COR molecules. This is the expected low energy cluster arrangement since the strength of intermolecular interactions increases with the number of carbon nearest neighbours [267].

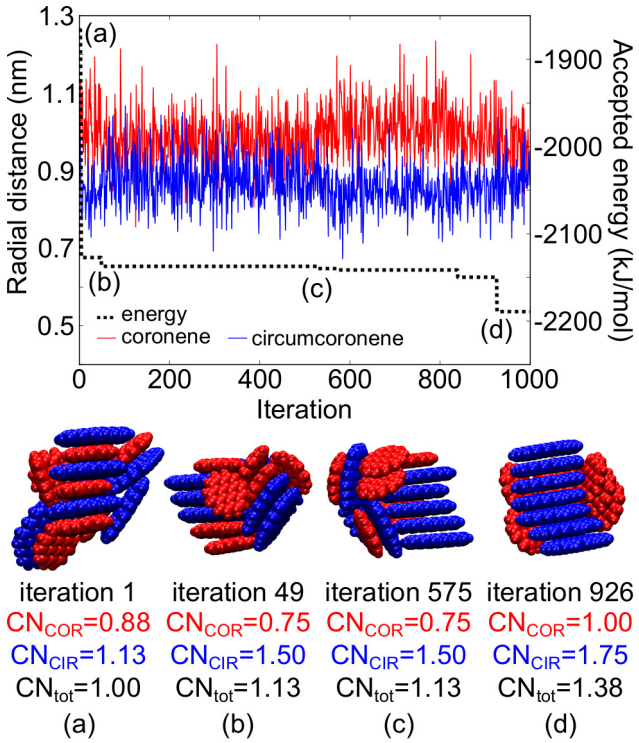


Fig. 6.4 Accepted energies and molecular radial distances as a function of iteration for the CIR_8COR_8 cluster. Cluster snapshots and corresponding molecular coordination numbers are shown for key minimum energy configurations. CIR molecules are shown in blue and COR molecules are shown in red.

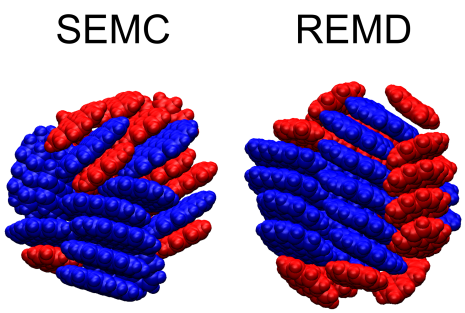
A partitioning of molecule types with the cluster is also seen quantitatively in the movement of the two molecule types, expressed as radial distances in Figure 6.4. Radial distances, measured as the average distance to the cluster centre for each molecule type during the course of the simulation, provide an indication of the separation of different molecule sizes within the cluster. The translational and rotational rearrangement steps in the optimisation procedure allow the molecule radial distances to vary considerably, but since the current lowest energy state is always used as the starting configuration for the rearrangement and minimisation of the next iteration, the divergence of the molecule radial distances highlights that the low energy structures have a configuration in which the larger molecules are closer

to the cluster centre than the smaller molecules. This molecule size partitioning agrees with recent experimental work on the molecule size distribution within carbonaceous soot particles collected from a flame [41].

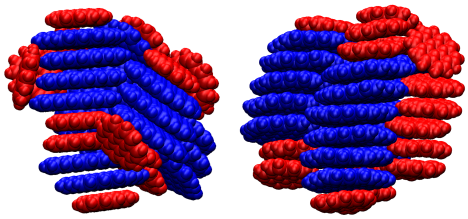
To provide an indication of the extent of stacking within each cluster, coordination numbers (*CN*) are calculated for the total system and each molecule type. The *CN* provides a simple indication of the number of nearest neighbours for each molecule type within the cluster. For the CIR_8COR_8 cluster, the degree of ordering of the cluster, as measured by CN_{total}, increases as the energy decreases. This is dominated by an increase of the close interactions and stacking by the CIR molecules, seen by an increase in CN_{CIR}. This is clearly indicated by the formation of a CIR column, moving from stacks mixed with COR at higher energies to CIR dimers and trimers to a central homogeneous CIR stack at the lowest energy. The interactions of the smaller COR molecules are secondary, showing an increase in stacking to form dimers and trimers as the energy decreases but only around the dominant CIR stack. In all simulations leading to the lowest energy configurations, low energy clusters are well stacked and a partitioning of the molecules is observed as the iterations progress, with the larger molecules found closer to the cluster centre than the smaller molecules, as seen here. In this way, a decrease in energy is seen to increase the stacking order of the molecules by causing more compact stacks to form and the two molecule types to partition. These trends are also seen within the OVA_8PYR_8 and larger heterogeneous clusters, and the results are recorded in Table 6.2.

Large PAH clusters compared to REMD

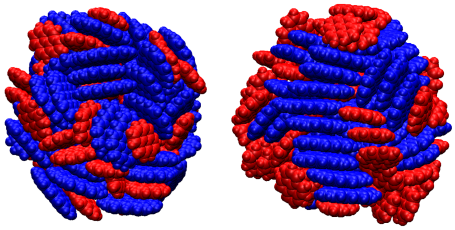
Minimum energy configurations of the $CIR_{16}COR_{16}$, $OVA_{16}PYR_{16}$, $CIR_{50}COR_{50}$, and $OVA_{50}PYR_{50}$ clusters are evaluated since they can be directly compared to previous REMD simulation results (in Chapter 4). Figure 6.5 shows the lowest energy structures for these large heterogeneous clusters, alongside the clusters determined by REMD. As with the 16 molecule system described above, the SEMC method finds low energy clusters in which the PAHs form stacks with the larger molecules within the cluster core. The smaller molecules are predominantly situated on the ends and sides of these stacks, although there is some mixing within the larger stacks. The clusters exhibit parallel stacks in a herringbone-like motif, which has been identified as the most stable configuration for large homogeneous clusters [267] and is similar to the bulk PAH crystal structure [184]. A parallel molecular columnar structure has also been observed experimentally in similarly sized PAH clusters [351].



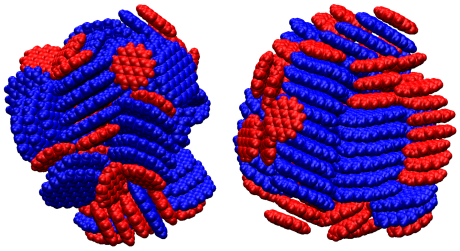
(a) $OVA_{16}PYR_{16}$



(b) $CIR_{16}COR_{16}$



(c) $OVA_{50}PYR_{50}$



(d) $CIR_{50}COR_{50}$

Fig. 6.5 Snapshots of final cluster configurations from SEMC simulations (shown on the left for each cluster pair) compared with REMD simulation results (shown on the right) for heterogeneous PAH clusters of different sizes. Larger molecules (CIR, OVA) are coloured blue and smaller molecules (COR, PYR) are red.

Compared to the CIR_xCOR_x clusters, the OVA_xPYR_x clusters present less structured motifs in which smaller stacks cause a weaker partitioning between the molecule types. This may be due to the smaller size and more elongated shape of the OVA and PYR molecules

compared to the CIR and COR molecules, which promotes electrostatically-stabilised T-shaped molecular configurations over the sandwich-like π-π interactions that form stacks. This is in agreement with previous work on homogeneous clusters that found large PAHs tend to adopt a herringbone-like structure more readily than small or linear PAHs [323, 317]. The smaller degree of stacking seen in clusters determined using the SEMC method may also be in part due to the use of a spherical volume in the encapsulating step required to prevent molecular overlap in the SEMC method. Sphere-containing clusters do not form stacked structures and thus this rearrangement step may introduce a bias unfavourable for planar molecules due to a misalignment between molecules. It is also seen that large 100 molecule clusters are less well-stacked than smaller clusters using both the SEMC and REMD methods, which suggests that the simulation length scales are not able to determine the global minimum energy configurations. However these results are still useful since the same cluster size trends are seen in experimental images of combustion-produced particles [107], indicating that processes creating these PAH clusters may not produce particles in a global minimum energy arrangement. It is well-known that many systems are strongly influenced by local minima and other features of the PES beyond the global minimum itself [345]. Thus low energy arrangements determined by the SEMC method for clusters containing a large number of PAHs may not correspond to the global minimum configuration but still provide useful information to relevant low energy systems.

A quantitative evaluation of these clusters, including the energies, radial distances, and *CN* values of the minimum energy configurations, shows good agreement between the SEMC and REMD methods, as recorded in Table 6.2. The SEMC method obtains minimum energy clusters with a 5% mean difference in intermolecular energy compared to REMD results for all clusters considered. Note that the intermolecular energies of heterogeneous PAH clusters are determined from an energy minimisation after the simulations in order to allow comparison between the two modelling methods. Similarly, the radial distances determined by the SEMC method are generally accurate to within 10% of the REMD results, ranging from 3–15% for the different cluster compositions and sizes. The molecular *CN* values and system CN_{total} values determined by the SEMC method present 20% and 10% mean differences, respectively. The largest differences, found in the CN_{small} values, can be identified by the tendency for the REMD produced clusters to form outer stacks made up of the smaller molecule type compared to the positioning of smaller molecules more independently around the cluster surface by the SEMC method. *CN* values provide an indication of molecular interactions and therefore the relative magnitudes of molecule types are more useful than a direct comparison. The high percent differences of the CN_{small} values are considered to be reasonable due to the large variation within this metric; an assessment

of the 10 lowest energy configurations determined by the SEMC method showed the percent standard deviation of CN_{small} values to be 25% averaged over all cluster types, in comparison to 2% for E_{int}, 3% for r_{large}, 4% for r_{small}, 6% for CN_{large}, and 7% for CN_{total} values.

Most importantly, the key structural motifs and molecule size partitioning trends observed from REMD results are evident in the minimum energy configurations determined by the SEMC method. Both the SEMC and REMD methods produce minimum energy clusters in which the molecules are well-stacked, as indicated by the high (> 1) CN_{total} values consistent across all cluster sizes and compositions. In all cases, the larger molecules (CIR and OVA) have higher CN values than the corresponding smaller molecules (COR and PYR), suggesting that they are less likely to be found on the surface of the cluster. The CN_{large} values are above one, corresponding to each molecule experiencing close interaction with two other molecules, indicating a stacked structure. The CN_{small} values are smaller, corresponding to surface molecules only located near one other molecule. Likewise, the r_{small} values are nearly always larger than the r_{large} values, describing clusters in which the smaller molecules are found further away from the cluster centre than large molecules. This agreement between the two methods provides an independent confirmation of the REMD-derived conclusion that low energy heterogeneous PAH clusters favour a partitioning of molecule sizes with the larger PAHs in the cluster centre and thus also confirms that the position potential implemented in the REMD simulations to prevent evaporation in the high energy replicas did not considerably affect the results observed in the low energy replicas.

Computational expense is reported in Table 6.2 to provide an approximate cost comparison between the SEMC and REMD simulation methods. This highlights the chief strength of the SEMC method since it is able to produce low energy structures using significantly less computational power than REMD, requiring only 8–13% of the computer time for the clusters studied. The SEMC method provides a means of determining the low energy configurations of homogeneous and heterogeneous aromatic clusters using significantly fewer computational resources than previous methods. This allows the simulation of aromatic clusters without the use of extensive supercomputing facilities, as required with REMD.

6.3.3 Evaluation of complex PAH clusters

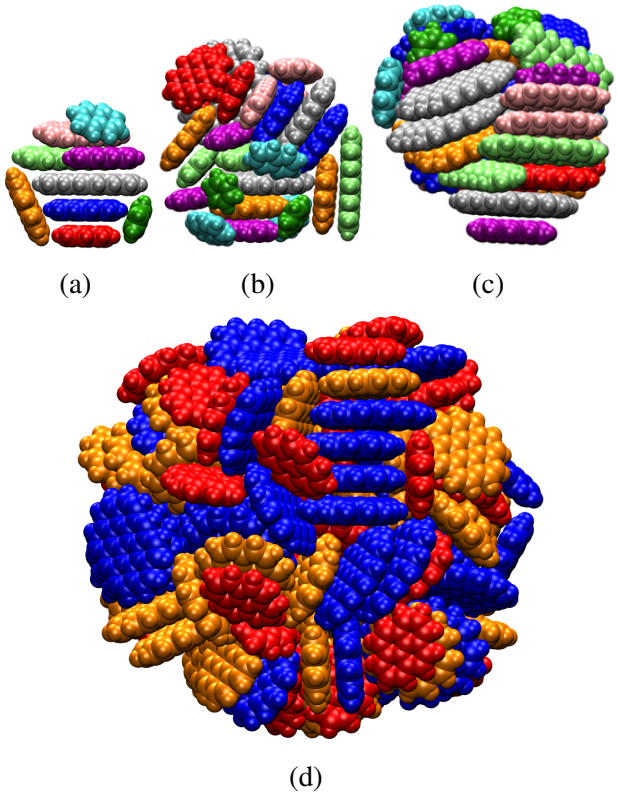
Two additional systems are considered to further illustrate the potential of the SEMC method: clusters containing nine different molecule types and a cluster containing 150 molecules of three different molecule types.

Table 6.2 Final intermolecular energy (E_{int}, kJ/mol), average radial distance (r, nm), coordination number (CN), and computation run time (CPU kilohours) of heterogeneous cluster systems from SEMC and REMD simulations. Values are obtained from post-simulation minimised clusters to allow comparison between the SEMC and REMD methods. The subscripts refer to the large (CIR, OVA) and small (COR, PYR) molecule types within each cluster as well as the total system.

Cluster property	Simulation method	CIR_8 COR_8	CIR_{16} COR_{16}	CIR_{50} COR_{50}	OVA_8 PYR_8	OVA_{16} PYR_{16}	OVA_{50} PYR_{50}
E_{int}	SEMC	-2220	-4650	-16100	-1320	-2920	-9420
	REMD	–	-4870	-15800	–	-2900	-10900
	% difference	–	5	2	–	1	15
r_{large}	SEMC	0.82	1.05	1.76	0.79	0.97	1.55
	REMD	–	1.02	1.57	–	0.89	1.38
	% difference	–	3	11	–	9	12
r_{small}	SEMC	1.02	1.35	1.86	0.86	1.07	1.49
	REMD	–	1.42	2.11	–	1.23	1.74
	% difference	–	5	13	–	14	15
CN_{large}	SEMC	2.0	1.9	1.7	1.6	1.6	1.4
	REMD	–	2.0	2.0	–	2.0	1.8
	% difference	–	5	16	–	22	25
CN_{small}	SEMC	1.0	0.81	0.68	0.88	0.56	0.52
	REMD	–	1.1	1.1	–	0.56	0.68
	% difference	–	30	47	–	0	27
CN_{total}	SEMC	1.6	1.5	1.5	1.4	1.4	1.2
	REMD	–	1.7	1.8	–	1.5	1.4
	% difference	–	13	18	–	7	15
CPU khours	SEMC	0.12	0.43	4.48	0.04	0.24	2.10
	REMD	–	5.32	56.2	–	1.80	23.0
	% difference	–	170	170	–	153	167

REMD requires significant computational resources, especially when considering heterogeneous clusters containing molecules of dissimilar sizes. This is because the difference in melting point of the constituent molecules demands a larger temperature range and thus an increase in the number of replicas required within the REMD method. SEMC simulations have no such dependency on the molecular melting points, allowing this method to consider a diverse set of heterogeneous clusters at a significantly lower cost than REMD. Thus there is great potential to use the SEMC method for highly inhomogeneous systems, which are

believed to be representative of particles found in real environments [320, 292]. As an example, Figure 6.6(a-c) shows the minimum energy arrangements of clusters containing 9 molecule types at three different cluster sizes. These results show similar behaviour to the binary heterogeneous clusters discussed previously, with stacked columnar structures dominating. The larger molecules also reside in the centre of these stacks and the smaller molecules reside around the outside, engaging in T-shaped interactions.



(a) (b) (c)

(d)

Fig. 6.6 Complex PAH clusters evaluated by the SEMC method. Minimimum energy clusters containing 9 aromatic molecules of varying sizes: benzene (dark green), naphthalene (cyan), anthracene (purple), tetracene (pink), pentacene (light green), pyrene (red), coronene (orange), ovalene (blue), circumcoronene (silver), denoted as $BEN_xNAP_xANT_xTET_xPEN_xPYR_xCOR_xOVA_xCIR_x$, are shown containing (a) one of each molecule type, $x = 1$; (b) three of each molecule type, $x = 3$; and (c) five of each molecule type, $x = 5$. The large $OVA_{50}COR_{50}PYR_{50}$ cluster is shown in (d).

A large cluster containing 150 molecules of OVA, COR, and PYR is evaluated to show the potential of the SEMC method to study systems above the reasonable size limit using the REMD method. The minimum energy structure of this heterogeneous cluster is shown in Figure 6.6(d). The molecular *CN* values, $CN_{OVA} = 1.22$, $CN_{COR} = 0.92$, and $CN_{PYR} = 0.54$,

indicate that the larger molecules more readily form stacks within the centre of the cluster than the smaller molecules, a result consistent with previous cluster structures.

Further work should be done to benchmark and improve the SEMC method. For example, the degree of stacking of the molecules may be tuned using the convergence criteria, defined as the maximum force less than a specified limit, however this may increase the required simulation time and one can also expect convergence problems due to accumulation of numerical errors. In addition, the molecular isoPAHAP minimisation is the slowest step in the SEMC process and it is observed that a significant amount of computational time is spent at the end for a small decrease in system energy, as an outcome of relatively negligible modification of the molecular configuration. There is therefore potential to increase the overall speed of the SEMC process by implementing a less strict convergence criteria for the molecular minimisation.

6.4 Conclusions

The Sphere Encapsulated Monte Carlo method is developed in order to efficiently determine minimum energy configurations of clusters containing aromatic molecules. The SEMC method builds on the framework of the basin-hopping method, in which a system is evaluated through a series of random perturbations and minimisations, through the addition of an intermediate sphere encapsulation of the molecules. The SEMC method is able to provide low energy cluster configurations of homogeneous coronene clusters in agreement with existing global minimisation methods.

The arrangements of PAHs in heterogeneous clusters are explored using the SEMC method. Binary PAH clusters containing up to 100 molecules present a parallel stacks motif, similar to the herringbone structure seen in bulk PAHs. The maximisation of carbon-carbon interactions dominates the arrangements formed, leading to smaller PAHs capping the ends of otherwise homogeneous stacks of larger PAHs. This agrees with previous work on small heterogeneous clusters and replica exchange molecular dynamics simulations. Minimum cluster energies are accurate to within 5% of REMD results using only a tenth of the computing resources. The SEMC method is also used to explore several complex PAH clusters beyond the scope of existing computational methods. In addition to its ability to determine good system minima with a fraction of the computational expense used previously, the SEMC method does not require additional constraints or large temperature ranges as required within REMD, and could be useful for generating minima to determine further system information such as transition state paths. The SEMC method provides a novel method of implementing a stochastic method

to determine low energy arrangements of aromatic clusters with significantly fewer resources than existing methods, useful for evaluating nanostructures present in PAH environments such as combustion particulates and interstellar dust.

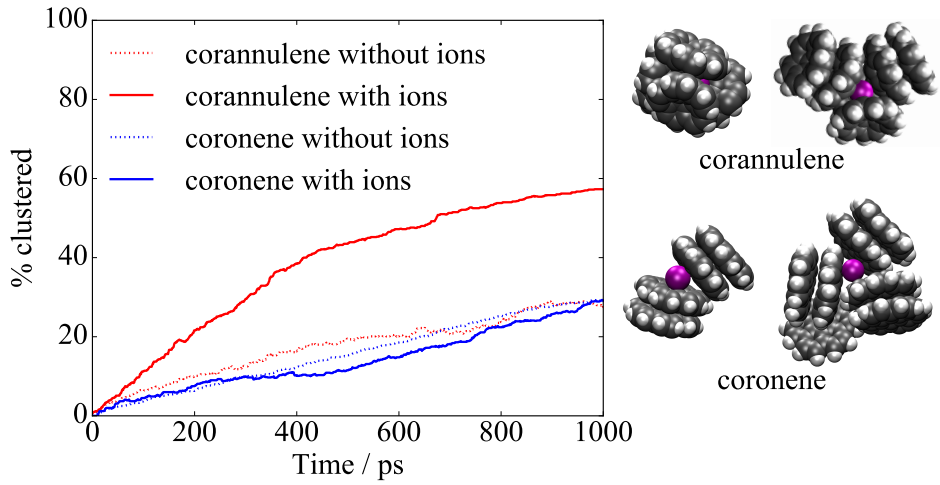
Chapter 7

A new potential for curved PAHs and its application to particle self-assembly

A potential able to capture the properties and interactions of curved polycyclic aromatic hydrocarbons (cPAHs) was developed and used to investigate the nucleation behaviour and structure of nascent soot particles. The flexoelectric charge polarisation of cPAHs caused by pentagon integration was included through the introduction of off-site virtual atoms, and enhanced dispersion interaction parameters were fitted. The electric polarisation and intermolecular interactions of cPAHs were accurately reproduced compared to ab initio calculations. This potential was used within molecular dynamics simulations to examine the homogeneous and heterogeneous nucleation behaviour of the cPAH corannulene and planar PAH coronene across a range of temperatures relevant to combustion. The enhanced interactions between cPAHs and potassium ions resulted in significant and rapid nucleation of stable clusters compared to all other systems, highlighting their importance in soot nucleation. In addition, the resulting cPAH clusters present morphologies distinct from the stacked planar PAH clusters.

Highlights:

- A new intermolecular potential, curPAHIP, is developed for curved polycyclic aromatic hydrocarbons.
- Molecular dynamics simulations show the clustering behaviour of corannulene and coronene with and without potassium ions across a range of temperatures.
- Systems containing corannulene and ions show significant clustering compared to all others, highlighting the importance of interactions between polar curved aromatics and ions.
- The morphologies of clusters containing curved PAHs show a "flower" motif in contrast to the "propeller" and stacked motifs of planar PAHs.



The work presented in this chapter was published as a paper in Combustion Science and Technology by Bowal et al. [44].

7.1 Introduction

The formation of soot particles is a complex chemical and physical process that involves gas phase chemistry of reactive precursor species and subsequent particle processes. The transition between gas phase molecules to solid particles is the least well understood aspect of soot formation. Proposed physical nucleation mechanisms suggest that dimerisation of polycyclic aromatic hydrocarbons (PAHs) plays a pivotal role, but the reversibility of this process, the high concentration of PAHs required, and the weak intermolecular interactions at high temperatures present difficulties [98, 352, 71]. Similarly, posited chemical inception mechanisms, which suggest stabilisation of nucleating particles by the formation of aliphatic links between aromatic molecules, do not fully describe all of the observed systems since they require a significant amount of hydrogen radicals and the predicted nascent particles have a much lower C/H ratio than is found in flames [352].

It was proposed that interactions between curved PAHs (cPAHs) and ions may play a significant role in soot formation [217, 216]. Curved aromatics, which differ from their planar counterparts through the integration of non-hexagonal rings into their predominately hexagonal arrangement, are found to comprise a significant portion of young soot particles [40, 350, 216]. This non-planar structure has been shown to produce significant polarity in otherwise non-polar molecules. For example, coronene and corannulene are both PAHs consisting of a single concentric arrangement of aromatic rings, although in contrast to the planar coronene molecule corannulene is curved due to a central pentagonal ring. Coronene is non-polar while corannulene has a dipole moment of 2.07 Debye [210], which is similar to that of water at 1.85 Debye [298] (although a better comparison is made with the local dipole moment at the pentagonal site of corannulene which is $\approx$1.63 Debye given in previous analysis [221]). This charge polarisation is primarily due to the strain-induced shift of the electron density from the concave to convex side of the curved molecule, known as the flexoelectric effect [221]. Experimental and computational work has determined that a representative nucleating soot molecule contains 15 rings, two of which are pentagons, and has a dipole moment of 5.32 Debye [216].

The unique electrical properties of curved aromatics allow them to interact strongly with ions, which has been recently explored for its potential in supramolecular and materials chemistry applications [109, 367]. Density functional theory (DFT) calculations have shown that the binding energy between a typical curved soot molecule and the chemi-ion $C_3H_3^+$ is 150–170 kJ/mol, suggesting that these interactions are strong enough to stabilise a small cluster of 1–1.5 nm at flame nucleating temperatures (1300–1500 K) [216, 124]. Experimental evidence

showing increased physical clustering with cations also supports an ionic mechanism of soot particle formation [53, 54]. Ion-dipole, dipole-dipole, and dipole-induced-dipole interactions between cPAHs and ions are long-ranged, with the former scaling with the reciprocal of the separation distance squared, and the other two as the inverse separation cubed and to the sixth power respectively. These long-range interactions are likely to be especially significant in combustion systems where species are present in low concentrations. Previous work has provided insight into these molecules and their interactions using electron microscopy and electronic structure calculations [216], but it remains unknown what effect these enhanced interactions have in a large system containing many cPAHs and ions. In order to develop more detailed understanding of the role of cPAHs and ions in the formation and structure of nascent soot particles, the simulation of dynamic systems is required to investigate the temperature dependent behaviour of cPAHs and ions and the morphology of resulting nanoparticles.

Molecular dynamics uses classical mechanics to study the movement and behaviour of molecular systems over time and provides valuable information about molecular interactions and arrangements of nascent soot particles [327, 64, 63, 165, 132, 71, 102]. These simulations require an accurate description of the interactions between constituent components, typically provided by an atom-atom potential in which the interaction energy at each point in time is calculated as a sum of pairwise interactions between atoms. An isotropic PAH potential (isoPAHAP) was developed using high accuracy intermolecular interaction energies from symmetry-adapted perturbation theory based on density functional theory (SAPT(DFT)) [232, 154, 259] to provide an accurate description of intermolecular interactions of PAHs [324, 327]. This potential was an improvement from previous general hydrocarbon potentials, such as the Lennard-Jones [334] and Williams' [361, 362] (W99) potentials, particularly at the T-shaped dimer geometries [324, 252].

The isoPAHAP potential uses atom-centred, molecule-specific point charge models derived from a transferable distributed multipole model involving both charges and quadrupoles to describe the electrostatic properties of PAHs [325]. This atom-centred charge-only representation is suitable for moderately sized planar PAH molecules for which the electrostatic interaction dominantly arises from the terminal groups where it can be modelled using atom-centred charges. In such systems there are no substantial local dipole terms within the PAH and the contributions from the quadrupolar terms are relatively minor compared to the dispersion and exchange-repulsion contributions. Consequently this model has been used successfully in a number of studies of PAH systems [64, 63, 132, 252]. However in cPAH molecules, strain at the carbon sites results in charge polarisation through the flexoelectric effect, and this leads to the formation of local dipole moments at the carbon atoms in the interior of the molecule. These dipole moments cannot be captured by the atom-centred

charge models used in the isoPAHAP model. For example, the isoPAHAP charge model incorrectly models the electrostatic potential around the curved corannulene molecule, as shown in Figure 7.2a. The dipole moment is found to be 1.73 Debye, significantly reduced from the calculated and experimentally determined values of 2.07 Debye [221]. An improved description of the electrostatic potential is provided by the distributed multipole expansion, which describes the charge distribution of a molecule using atom-centred multipole functions [310, 309], but this is computationally expensive to use and not widely implemented in molecular dynamics software. In order to examine systems containing curved aromatics, a new potential needs to be developed that describes the flexoelectric effects in a computationally efficient and readily usable implementation.

This chapter aims to study the role of cPAHs and ions in stabilising nascent soot particles by investigating their clustering behaviour using an accurate intermolecular potential. In Section 7.2 an intermolecular potential (curPAHIP) is developed that is able to capture the flexoelectric effect and enhanced interactions of cPAHs and ions. Simple interaction energies for the small cPAH corannulene are compared with *ab initio* calculations. In Section 7.3, the curPAHIP potential is used within molecular dynamics simulations to provide information on homogeneous and ion-induced heterogeneous nucleation of corannulene, in comparison to analogous systems containing the planar coronene molecule. Clustering behaviour and resulting morphologies are explored across a range of temperatures relevant to soot formation. These results provide insight into how the interactions between cPAHs and ions affect the formation and structure of soot nanoparticles and provide valuable information towards expanding current soot models.

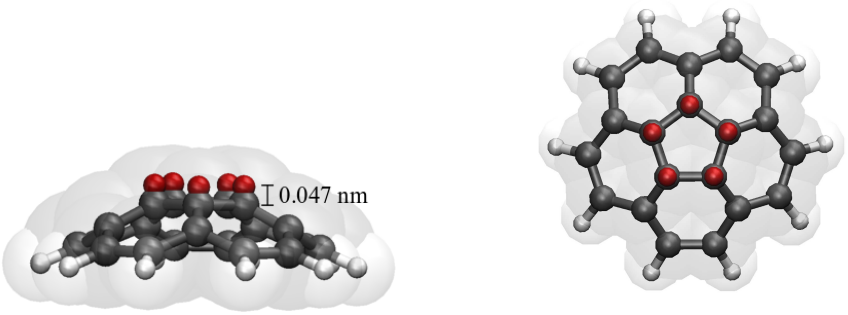
7.2 Developing a potential for curved PAHs

7.2.1 Determining the electrostatic potential

The isotropic all-atom isoPAHAP potential is used as the foundation for developing a potential able to describe curved PAHs. The form of the isoPAHAP potential, and also corresponding to the curPAHIP potential, is described in Equation 3.3.

In this chapter the cPAH corannulene ($C_{20}H_{10}$), consisting of five hexagonal rings surrounding a pentagonal ring, is used as a representative cPAH. Although it is likely too small to contribute significantly to soot nucleation at flame temperatures [216], it provides a good starting point for the development of a cPAH potential with a distinct benefit of having been studied in previous *ab initio* studies. To incorporate the charge distribution created

by curving a PAH, off-site point charges are added to the molecule description in a method similar to the multiple site models of water [27, 178]. These mass-less virtual sites are fixed at a distance of 0.047 nm directly above each of the pentagonal carbons on the convex surface of the cPAH, as seen in Figure 7.1. This distance and the corresponding point charge values, found in Appendix B, were optimised to match DFT calculations which provide a good description of the electric potential around corannulene [221]. This was done using the MULFIT programme, which fits atomic charges through a systematic reduction of the distributed multipole expansion [106].



(a) Side view (b) Top view

Fig. 7.1 Geometry of the curved PAH corannulene ($C_{20}H_{10}$) with added off-site charges fixed above the pentagonal carbons. Carbon atoms are shown in grey, hydrogen atoms are shown in white, and off-site charges are shown in red.

This modified molecule description allows the dipole moment of corannulene to increase from 1.73 Debye (using fitted atom-centred charges [221]) to a value of 2.14 Debye, in reasonable agreement ($< 4\%$ difference) with the known dipole moment of corannulene. Figure 7.2b shows that the electrostatic potential around corannulene after including the off-site point charge description is in good agreement with DFT results.

The use of the OPLS-AA forcefield for the intramolecular forces produces a geometry in good agreement to the corannulene structure determined by X-ray crystallography [255] and DFT calculations [221]. Corannulene is known to invert at room temperature [295], however molecules are considered rigid in these simulations since an inversion would require a coupling of the atomic charges with the geometry change. This inversion barrier rapidly increases as rings are added and therefore this model is appropriate for studying the interaction of ions and large curved arenes, which are observed in the flame [216].

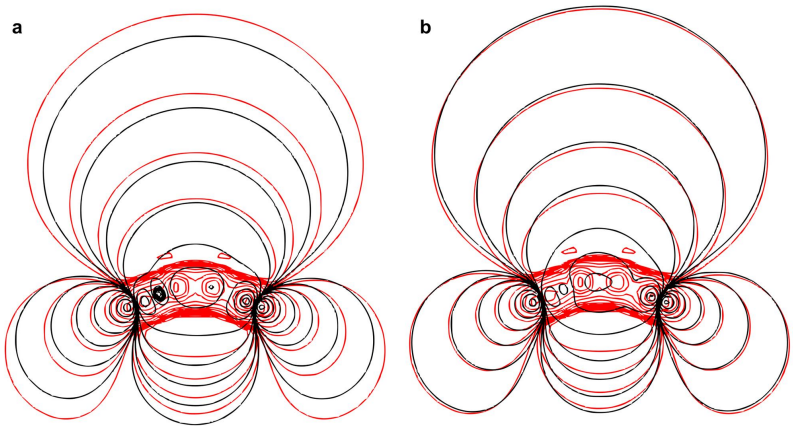


Fig. 7.2 Comparison of the electrostatic potential of corannulene calculated using DFT (red) and a potential description for implementation in molecular simulations (black). (a) shows the isoPAHAP potential using atom-centred charges. (b) shows the curPAHIP potential in which the electrostatic potential more closely matches the DFT results by using off-site charges.

7.2.2 Dimer binding energies

The binding energy of a corannulene dimer was assessed as a function of the dimer separation distance using several potential descriptions and the results are shown in Figure 7.3. The bowl-to-bowl "sandwich" configuration was considered since this is likely to be a low energy dimer configuration and SAPT(DFT) interaction energies are available [49]. All evaluations of the intermolecular potentials were performed using the ORIENT programme [311].

From Figure 7.3 it is evident that the interaction energy is significantly underestimated using the isoPAHAP potential. Some disagreement could be expected because, as discussed previously, the atom-centred charges used in the isoPAHAP potential underestimate the flexoelectric effect. However changing the electrostatic model in the isoPAHAP potential to a full multipole model derived specifically for corannulene or adding off-site charges do not cause a significant improvement in the interaction energies and the error remains large at the energy minimum separation. This is because the presence of pentagonal rings increases the dispersion interaction of the molecule by about 20%. This increase is known to occur for curved carbon nanostructures [126].

We have addressed this difference through a re-optimisation of the relevant isoPAHAP parameters, specifically the dispersion coefficients ($C_{6,ab}$) and shape function (ρ_{ab}), to result in the curPAHIP model. The parameters were simultaneously refitted using SAPT(DFT) energies, which were weighted as in Totton et al. [324] in order to favour more negative

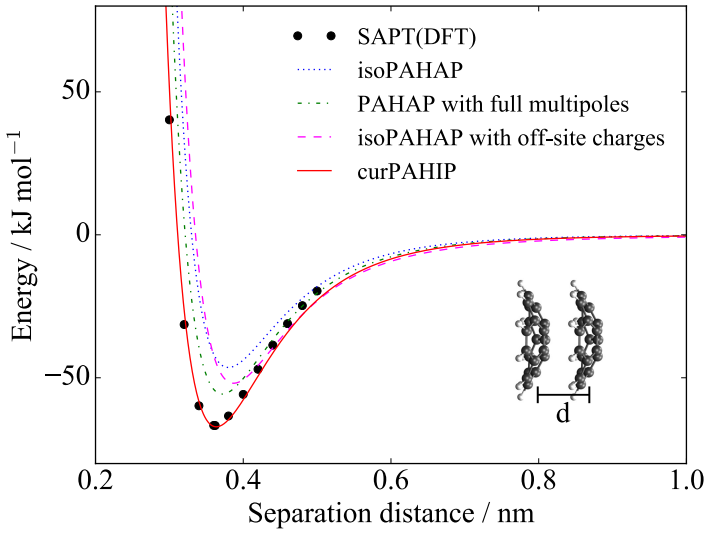


Fig. 7.3 Interaction energy versus separation distance for a corannulene dimer determined from SAPT(DFT) calculations [49], the isoPAHAP potential, the PAHAP potential with full multipoles, the isoPAHAP potential with off-site charges, and the isoPAHAP potential with off-site charges and re-fitted dispersion parameters (curPAHIP).

energies and thus ensure an accurate potential well. Harmonic constraints were used to prevent the parameters of curPAHIP from deviating too far from the corresponding isoPAHAP values. The resulting parameters are shown in Table 7.1 (SI units provided in Appendix B), and provide a weighted root mean square residual energy of 1.30 kJ/mol. Following Misquitta and Stone [233] the dispersion damping term β is not fitted but is defined through the molecular vertical ionisation energy I as $\beta = 2(2I)^{1/2}$. $\beta = 1.50$ a.u. is obtained using the value of I for corannulene [288].

Table 7.1 Parameters of curPAHIP in a.u.

Atom pair	ρ	α	C_6
C C	5.6563	1.8783	30.282
C H	4.9320	1.7560	12.604
H H	4.1187	1.4043	5.2179

As can be seen in Figure 7.3, the curPAHIP potential is able to capture the interaction behaviour of corannulene using off-site point charges and optimised dispersion parameters.

7.2.3 K^+ binding energies

SAPT(DFT) results of corannulene dimers show that induction energy, defined as the attractive energy due to the electron density of a molecule distorting in response to the electric

field of another nearby molecule, is a very small component of the interaction, contributing to only 4% of the attractive energy [49]. The curPAHIP potential, like isoPAHAP, does not model the induction energy (a combination of the classical polarisation and charge-transfer energies) as separate term, but instead absorbs the effects of the second-order induction into the dispersion and exchange-repulsion parameters. This is reasonable when the induction energy is a relatively small component of the total interaction energy, as it is in the interactions of both planar and curved PAH molecules. However the induction plays a significant role when ion interactions are involved, with energy contributions equal to or greater than those from the electrostatic component [331, 215, 293]. Incorporating a physical term that includes dynamic polarisation and charge-transfer into atomic models comes at a very high computational cost due to the need for local coordinate systems for each atomic multipole, iterative solving of mutually induced dipoles, and the increased complexity of intra- and intermolecular terms. Instead it is sometimes possible to include polarisation effects indirectly within non-polarisable potentials with fixed charges by computing a set of polarised point charges obtained by treating the surrounding molecules in an average manner, typically via the introduction of a dielectric with suitable dielectric constant. Models obtained in this manner can be expected to be limited in that they are only applicable to the system they are designed for and the effective charge model is valid only for large-scale homogeneous systems.

A similar strategy was employed here and an effective interaction model was developed for a system containing potassium ions, K^+, and corannulene molecules. An effective charge model may not be appropriate for a heterogeneous system such as this since the molecular environment, and thus the polarisation, is not fixed but varies in space. Therefore an effective interaction model was developed by altering the exchange-repulsion and dispersion terms in a Lennard-Jones potential form. This was done in a fitting procedure initialised from the OPLS-AA parameters for K^+ [172]. Ensuring accurate binding energies was prioritised and so the K^+ parameters were fit to the maximum binding energy of corannulene and K^+ determined from hybrid DFT with an empirical dispersion correction (B97D/cc-pVTZ), which has been found to accurately describe binding energies of Na^+-benzene interaction to within 1.3 kJ/mol compared with CCSD(T)/CBS calculations [240]. The resulting parameters, $\sigma = 9.64$ and $\varepsilon = 3.43 \times 10^{-8}$ a.u. (SI units provided in Appendix B), provide an interaction energy of -87.7 kJ/mol. This agrees well with the binding energy of -88 kJ/mol determined from B97D. However the model results in a reduced equilibrium separation distance of 0.25 nm compared to 0.29 nm from DFT calculations. The underestimation of the separation provided by this fit is considered acceptable for the work discussed here since its potential

consequences, such as increasing steric effects and corannulene-corannulene repulsion within clusters, result in a conservative estimate of K^+-corannulene interactions and clustering.

7.3 Nucleation simulations

Molecular dynamics simulations were used to investigate the homogeneous and ion-induced heterogeneous nucleation of the small curved aromatic molecule corannulene ($C_{20}H_{10}$, 250 g/mol) and its planar counterpart coronene ($C_{24}H_{12}$, 300 g/mol). The clustering behaviour and resulting morphologies were explored over a range of flame relevant temperatures. As the smallest curved PAH, corannulene is often used as a model molecule for fullerene-like species [29]. Thus its molecular properties, such as boiling and melting points, are established and it provides a useful system for assessing the behaviour of cPAHs in flames. Corannulene has been found to be present in flames [194], however its clustering effectiveness is expected to be less significant than those for larger curved molecules prevalent in sooting environments due to its weaker interaction energy. Potassium is readily ionised in a flame and has been shown to impact soot nucleation through the production of higher particle numbers and smaller particle sizes [145, 82, 300]. Due to this known influence and the straightforward incorporation of the spherical potassium cation within an atomic model, K^+ provides a representative cation useful for exploring the interactions of cPAHs and ions in a large dynamic system. This system thus presents an exploratory study comparing the interactions of small curved and planar PAHs with and without cations.

7.3.1 Molecular dynamics methods

Molecular dynamics simulations were performed using the software GROMACS 5.1.4 [4] according to a similar methodology detailed in previous work [327]. One thousand PAH molecules (coronene or corannulene) with and without the same number of potassium ions were randomly placed in a periodic box. This ratio and large number of molecules ensured proper statistics of molecular interactions and cluster stability. A concentration of 2×10^{18} cm^{-3} allowed experimentally relevant processes to be observed in the simulation timescale, as described in Totton et al. [327]. A two-step minimisation process using the steepest descent and low-memory Broyden-Fletcher-Goldfarb-Shanno approaches removed any excess energy caused by the initial arrangement. Canonical simulations, defined by maintaining a constant number of atoms, system volume, and temperature, were then conducted for 1 ns with a timestep of 1 fs using a velocity Verlet integrator. A chain of 10 Nosé-Hoover

thermostats was used for temperature control and intermolecular cut-offs were set to 3.0 nm. Electrostatic interactions were calculated at long range using particle-mesh Ewald summation. Intramolecular forces were determined using the OPLS-AA force field for molecule bonds, angles, dihedral and improper dihedral angles [179]. The isoPAHAP potential was used to describe the coronene interactions while the newly developed curPAHIP intermolecular potential was used to describe corannulene interactions and enhanced Lennard-Jones parameters defined the interactions between K^+ and corannulene.

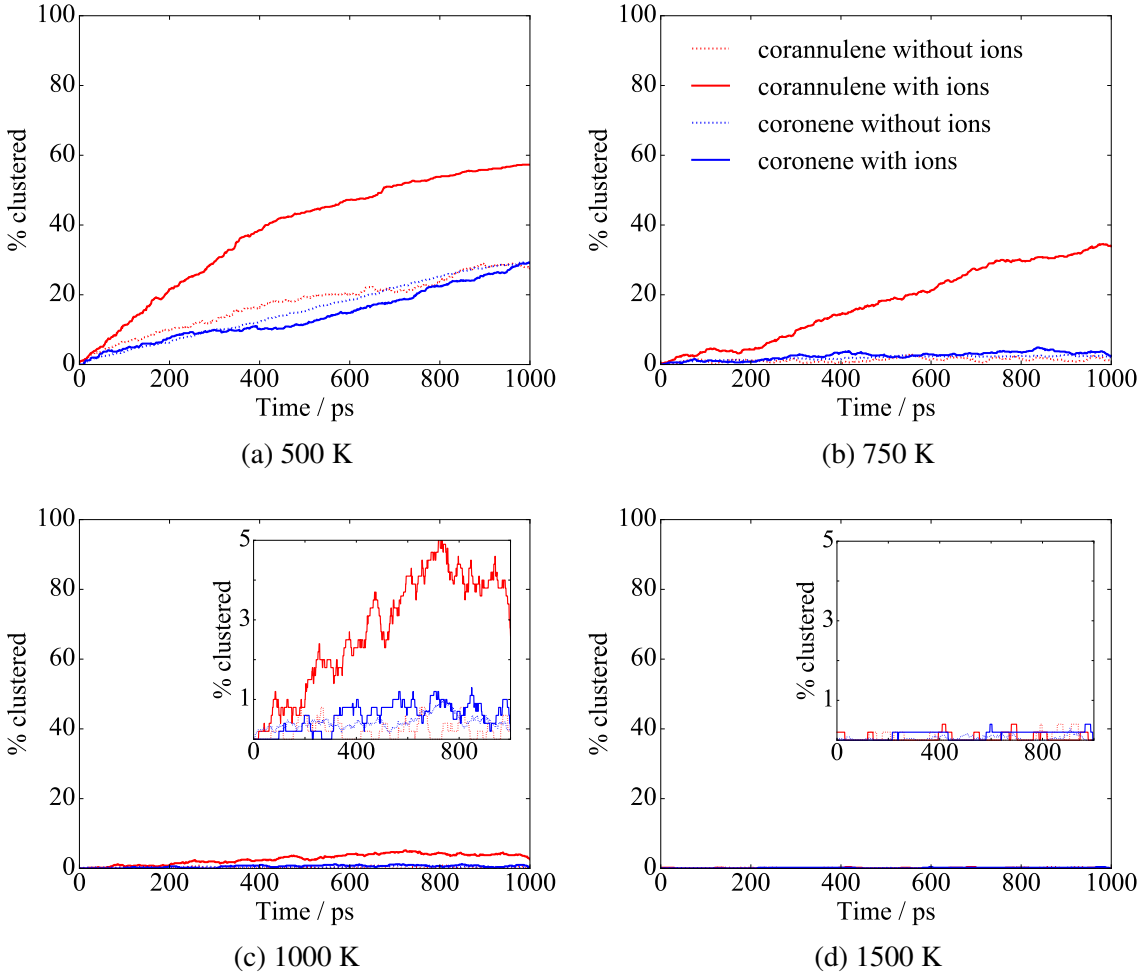
The clustering behaviour and resulting morphologies were explored at 500, 750, 1000, and 1500 K. Clusters were identified using a previous methodology that was shown to provide a suitable measure of stable clusters for PAHs ranging from 4 to 19 rings in size [327]. This means that PAH molecules must be within 1.2 nm of each other for at least 20 ps to be considered bound in a cluster. Collision efficiencies and lifetimes were calculated to provide an assessment of nucleation propensity and stability of formed clusters. It is important to note that in this analysis a cluster must contain at least two PAHs, i.e. a single PAH interacting with an ion is not considered to be a cluster. Finally, cluster sizes and morphologies were also discussed.

7.3.2 Clustering over time

Figure 7.4 shows the extent of clustering for each system over time. The degree of clustering decreases with increasing temperature such that for all systems clusters are formed at 500 K but no lasting clusters are formed at 1500 K. At 500, 750, and 1000 K the coronene, coronene with K^+, and corannulene systems all experience a similar extent of clustering. This similarity between homogeneous corannulene and coronene systems is unsurprising since they have very similar binding energies (-65.7 kJ/mol for a coronene dimer [257] and -65.9 kJ/mol for a corannulene dimer [49]). The clustering propensity of coronene molecules is unaffected by the inclusion of cations in the system, in agreement with electronic structure calculations that suggest the π-cation interactions are not strong enough to impact clustering of planar PAHs [65].

In contrast, corannulene molecules with K^+ ions experience significantly more clustering than all other systems, with an increase of about 50% above the other systems at 500 K, 95% above at 750 K, and 92% above at 1000 K. The degree of corannulene clustering with K^+ is in fact very similar to that of the much larger ovalene molecule ($C_{32}H_{14}$, 398 g/mol) in a homogeneous system [327]. Of particular interest is the fact that this enhanced clustering occurs to a significant degree even at 750 and 1000 K, which is above the sublimation point

of corannulene at 640 K [68]. The number of clusters formed as a function of the simulation time (not shown) shows the same trends as those seen when examining the percentage of molecules clustered over time, indicating that the increase in clustering is primarily due to the formation of new clusters instead of cluster growth.



(a) 500 K (b) 750 K

(c) 1000 K (d) 1500 K

Fig. 7.4 Extent of molecules clustered over time for corannulene (red) and coronene (blue) systems at (a) 500 K, (b) 750 K, (c) 1000 K, and (d) 1500 K. Solid lines represent systems with potassium ions present and dashed lines indicate systems containing no potassium ions. Inset figures provide a closer look at the clustering trends at high temperatures.

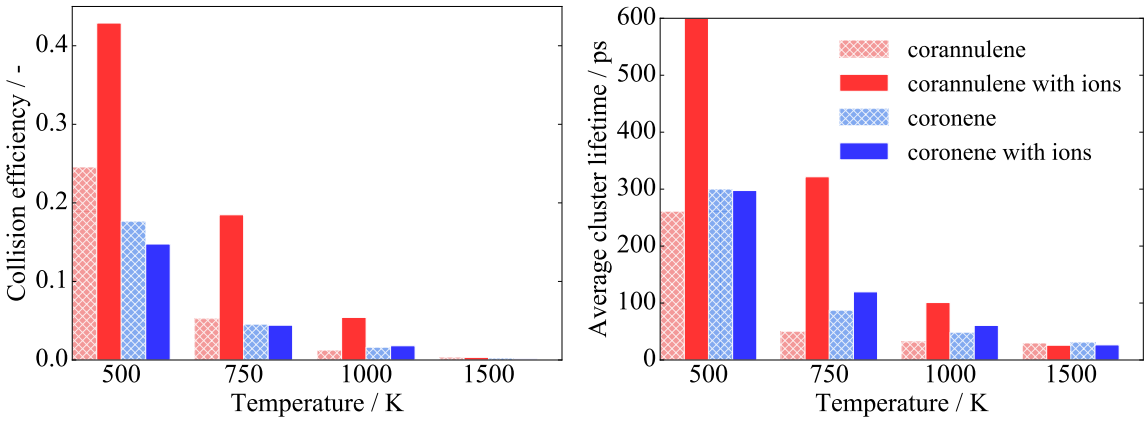
7.3.3 Collision efficiency

Further cluster statistics were evaluated to understand how these ion-cPAH interactions are able to increase the amount of clustering seen. Collision efficiency is defined as the number

of successful collisions over the total number of collisions and thus indicates the "stickiness" of the interacting components.

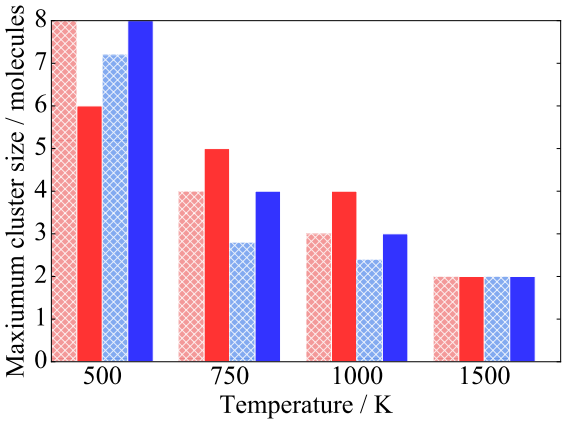
Collision efficiencies for all the systems and temperatures studied in this chapter are shown in Figure 7.5a. Collision efficiencies decrease significantly with temperature since the increased kinetic energy serves to both increase the total number of collisions and decrease the ability for the molecules to bind strongly. The collision efficiency is the highest for corannulene molecules with K^+ ions, with a value above that observed for significantly larger planar PAHs [327]. As seen in the degree of clustering over time, the inclusion of K^+ ions in the system containing coronene molecules has minimal effect on the cluster behaviour at all temperatures studied. This is due to the weak bonding between the planar PAHs and cations and indicates that the ions do not act as seeds around which the clusters of planar PAHs form. In the simulation trajectories, few coronene molecules were observed to interact with the potassium ions even at 500 K, whereas a significant portion of corannulene molecules are bound to a cation at that temperature.

Interestingly the nature of the high collision efficiency for corannulene systems containing K^+ ions is not the same across all temperatures studied. Above 1000 K the total number of collisions is similar between all the systems and so this high efficiency comes from a significantly higher number of successful collisions. This highlights the importance of the ion-cPAH interaction in promoting nucleation at high temperatures. The inverse is true at 500 K: the number of successful collisions is relatively consistent across systems but the total number of collision events is more than 50% lower for corannulene with ions. This may be because clustering occurs more rapidly in the corannulene and K^+ system. Due to its larger mass, the average velocity of a cluster is lower than that of a single molecule at a given temperature, so an increase in the number of clusters reduces movement within the system and results in fewer collision opportunities. In addition, at low temperatures repulsive interactions between ions and polar cPAHs may play a significant role and prevent the short-lived cluster formation seen when weak dispersive interactions dominate the system. This is supported by the average cluster lifetimes reported in Figure 7.5b, discussed in more detail later. At 750 K the number of successful collisions is approximately 275% higher and the number of total collisions is about 30% lower for corannulene with K^+ compared to the other systems. These results show the ability of ions to promote clustering efficiency of cPAHs across a wide range of temperatures.



(a) Collision efficiency

(b) Average cluster lifetime



(c) Maximum cluster size

Fig. 7.5 Cluster properties over 1 ns simulations containing planar or curved PAHs with and without K^+ ions. (a) provides collision efficiencies, (b) reports average cluster lifetimes, and (c) presents maximum cluster sizes. Systems containing corannulene are shown in red, systems containing coronene are shown in blue, and the hatching indicates systems without K^+ ions present.

7.3.4 Cluster lifetimes

Cluster lifetimes show the stability of clusters after they are formed, which is crucial in the understanding of soot formation. Average cluster lifetimes, shown in Figure 7.5b, decrease with temperature in all cases as expected. At low temperature (500 K), cluster lifetimes are long with some clusters lasting for nearly the entire simulation lifetime in all systems studied. At higher temperatures (750 K, 1000 K) the cluster lifetimes of corannulene molecules with K^+ ions are significantly higher than for other systems. At 1500 K, the average lifetimes of all systems studied are very low. Corannulene without ions exhibits slightly smaller cluster

lifetimes compared to homogeneous coronene, as expected according to its molecular weight. However the average cluster lifetimes are much higher for corannulene with K^+ than for all the other systems, with an increase of up to 375%. These results show that at temperatures below 1500 K, the presence of the ion is significant in promoting long cluster stabilisation for corannulene molecules.

7.3.5 Cluster sizes and morphologies

Figure 7.5c shows the maximum cluster size obtained over the simulations studied. The maximum cluster sizes decrease as the simulation temperature increases such that only occasional dimers are formed at 1500 K. Despite significantly higher collision efficiencies for systems containing corannulene and K^+ ions compared to the other systems studied, their maximum cluster sizes are relatively similar. This indicates that for the simulation length considered here, the presence of ions has a more significant role in nucleation compared to cluster growth. Figure 7.6 provides snapshots of the maximum clusters formed at each temperature within the systems containing corannulene and K^+ ions.

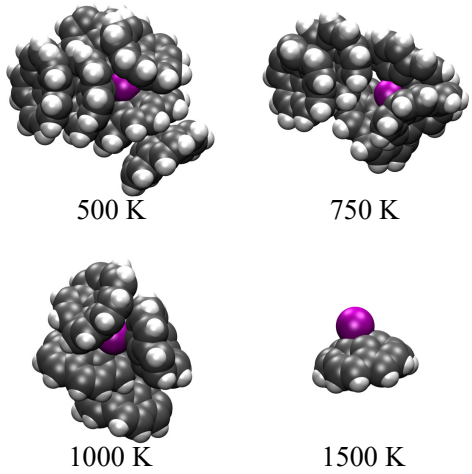


Fig. 7.6 Maximum clusters and complexes formed in simulations containing corannulene molecules and K^+ ions at 500, 750, 1000, and 1500 K. At the highest temperature (1500 K) dimers are rare and short-lived, so the more common and stable complex of K^+ and corannulene is shown here. Carbon atoms are shown in grey, hydrogen in white, and potassium in purple. Off-site virtual atoms are not shown.

The morphologies of clusters formed from cPAHs are significantly different than those made up of planar PAHs. Figure 7.7 shows cluster snapshots at 500 K for corannulene and coronene simulations with K^+ ions. These images indicate the typical and maximum cluster sizes over time and show observed cluster motifs.

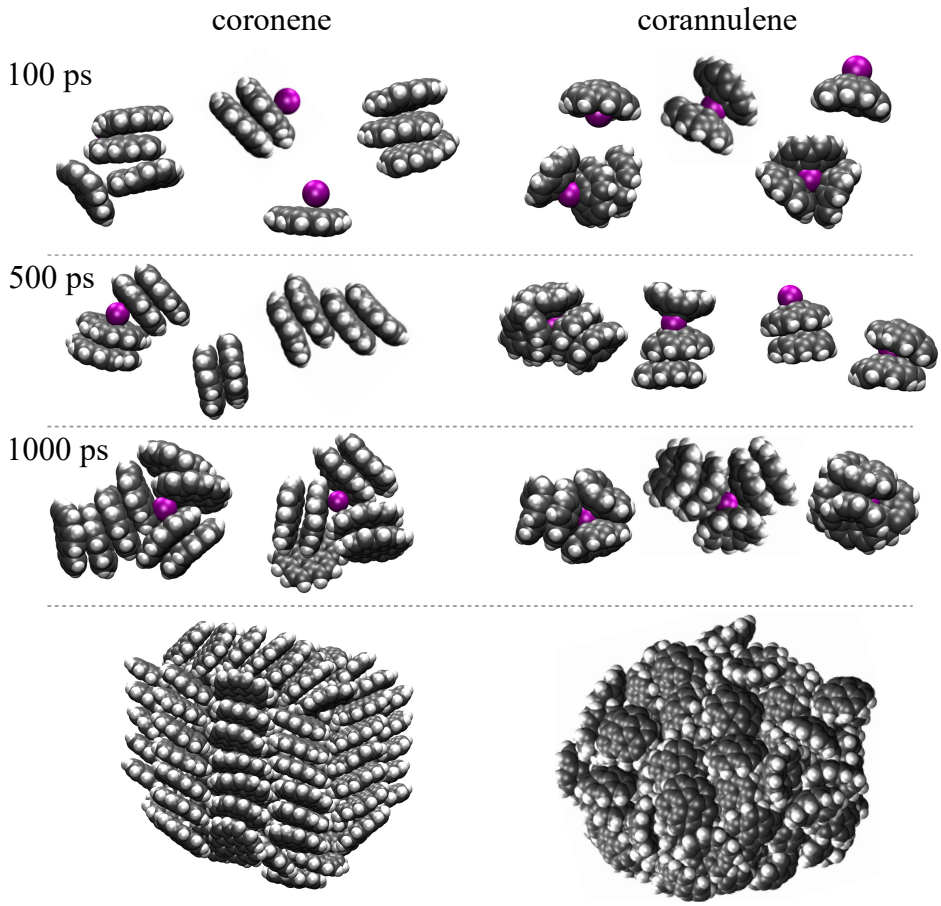


Fig. 7.7 Coronene and corannulene clusters formed at 500 K in the presence of K^+ ions at 100, 500, and 1000 ps. Expected arrangements of large clusters after an extended simulation without ions are shown for coronene [64] and corannulene molecules (bottom). Carbon atoms are shown in grey, hydrogen in white, and potassium in purple. Off-site virtual atoms are not shown.

Coronene molecules initially interact with K^+ ions in some cases although these interactions appear short lived and homogeneous clusters are more readily formed at 500 K. Homogeneous clusters primarily form single stacks, with some T-shaped interactions observed. Once a cluster grows to about four or five molecules, an ion is usually integrated into the structure. The planar PAHs cluster around the ion in a "propeller" motif, a staggered triangular arrangement which allows the cluster to be stabilised by cation-π and CH-π interactions. A maximum of three coronene molecules form the first solvation shell around the ion and dispersion interactions allow π-π stacking of further molecules.

In contrast to coronene, almost all corannulene dimers and clusters contain K^+ ions from very early in the simulation. The majority of K^+-corannulene interactions occur on the convex side of the curved molecule and only a few cases in which the K^+ interacts with the

concave surface are observed. This is in agreement with DFT results, which found that the K^+-corannulene complex is more stable when the cation is on the convex surface compared to the concave side [55]. The alignment of corannulene molecules around K^+ maximises electrostatic interactions. The first solvation shell around the ion consists of three or four molecules, with their convex surfaces pointing towards the ion - a "flower" motif. This arrangement allows for further interactions with additional molecules, held together by a combination of dispersion and electrostatic interactions, to form shifted stacks extending out from the ion. Up to four corannulene molecules are able to bind strongly with a cation in contrast to the three molecules found for small planar PAHs [65].

An additional simulation of corannulene without K^+ ions was conducted for 10 ns at 500 K to provide a preliminary expected morphology of a large cluster of cPAHs for comparison with the known structure of large clusters of planar PAHs used in many previous molecular modelling studies investigating soot properties. Due to strong dispersive interactions, a tightly stacked structure is favoured for low energy clusters of planar PAHs, such as the coronene cluster containing 50 molecules shown in Figure 7.7 obtained from Chen et al. [64]. In contrast, the large corannulene cluster shown here presents a more amorphous arrangement containing many disordered stacks containing 2–4 molecules. The cPAHs are often aligned with curved bowls slightly shifted or arranged such that neighbouring molecules alternate in direction. These motifs are similar to those seen in studies examining the packing structure of solid state cPAHs [110]. π-π and electrostatic interactions promote the formation of small stacks, while the flexoelectric dipole enhances the CH-π interactions between the positively charged rim of one molecule and the negative π-surface of its neighbour which promote arrangements in which molecules are perpendicular to each other. The combination of these two interaction types serves to stabilise the three-dimensional structure of the corannulene cluster. This initial insight suggests that cPAHs promote a spherical cluster arrangement more similar to the concentric-ringed structure observed in microscopy images of soot [10] but further work needs to be done to characterise the structure of cPAH clusters.

7.4 Discussion

No significant clusters were formed at 1500 K, indicating that even with stronger ion-cPAH interactions compared to planar PAHs, corannulene cannot contribute substantially to soot particle nucleation at flame temperatures. However, corannulene molecules did cluster around cations above their melting point, which suggests that the interactions between cPAHs and ions promote the nucleation process at high temperatures. This effect is expected

to be increased with larger cPAHs and smaller chemi-ions. Larger cPAHs have stronger electrostatic and dispersion interactions due to their larger dipole moment and molecule size, and this may provide sufficient additional binding energy to stabilise a cluster at flame temperatures. In addition, the binding energy between a corannulene molecule and K^+ ion is weaker than with a smaller cation due to a decrease in induction, a larger distribution of charge over the ion, and steric repulsion, particularly between the concave surface and large K^+ ion [215]. This is especially significant because for small cations the binding energy is greater between the cation and the concave surface of a cPAH compared to the convex surface, and this effect increases as the size of the cPAH increases [55]. The fact that a small cation binds very readily to the concave surface of a larger cPAH is supported by the high interaction energy (-170 kJ/mol) of a 15-ringed cPAH with $C_3H_3^+$ [216]. This arrangement would perhaps result in a nucleus structure which more closely resembles the experimentally observed seeds of primary soot particles from which an onion-like structure forms.

The models used in this study are based on accurate *ab inito* data, but they have important shortcomings: first, the present version of the curPAHIP potential does not take into account the increased polarisation of the cPAHs in proximity with the cation. This could be accounted for using an explicitly polarisable model, but while such models can be developed, and indeed have been for smaller organic molecules [234], this development was not pursued here to the increased computational cost of polarisable models. Second, while the cation-cPAH model used here is reasonable, the use of DFT as the reference and the too-short minimum configuration obtained with the model both lead to uncertainties; in particular, the latter error would lead to an increased steric repulsion between the cPAHs directly interacting with the cation. It is quite possible that a more advanced model, and one more carefully fitted to accurate *ab initio* data, will lead to more binding in systems such as those studied here.

Nevertheless, this chapter clearly illustrates the importance of considering the enhanced interactions present in curved carbon systems and provides insight into how the interactions between curved carbon molecules and ions affect the formation and structure of particulates in flames. Curvature has also been observed in many carbon structures such as glassy carbon, activated carbon and carbon blacks [143, 142, 321] and it is anticipated that the curPAHIP potential will be a valuable tool for describing the flexoelectric effect present in materials relevant to many other applications such as gas storage, separations, and microelectronics.

7.5 Conclusions

Curved PAHs containing five-membered rings are present in many carbon materials, including soot. Modelling the interactions between these molecules presents a challenge due to the flexoelectric dipole moment and enhanced dispersion interactions caused by the curvature. A new potential, curPAHIP, was developed using the foundations of the previously developed isoPAHAP potential. In this potential, the increased polarity of the cPAH is addressed with a modified molecule description in which off-site point charges are added to the pentagonal carbon atoms. This method was shown to properly describe the electrostatic potential and flexoelectric dipole moment of the small cPAH corannulene. The dispersion and shape function potential parameters were re-optimised to match SAPT(DFT) energies and enhanced parameters were determined to include polarisation with K^+ ion interactions.

We have demonstrated the value of the curPAHIP potential by applying it to a system relevant to soot formation. The clustering behaviour of corannulene molecules with and without K^+ ions was investigated using molecular dynamics and compared to analogous systems of the planar coronene molecule from 500–1500 K. In comparing PAHs with similar sizes and symmetry, the effects of curvature and cations in the nucleation of PAHs were highlighted. The clustering behaviours of coronene, coronene with K^+, and corannulene systems are very similar for all properties examined. In comparison, enhanced interactions between the polar corannulene molecules and K^+ ions result in more rapid and abundant clustering, higher collision efficiencies, and longer cluster lifetimes. The morphologies of corannulene clusters present a concentric arrangement in which bowl-to-bowl and T-shaped interactions are stabilised by dispersive and electrostatic interactions. This is different than the highly stacked structure of planar PAHs and perhaps more representative of experimentally observed soot particle structure. This chapter extends the current understanding of the role of planar and curved PAHs and ions in soot particle formation and provides an intermolecular potential valuable for the study of systems containing curved carbon.

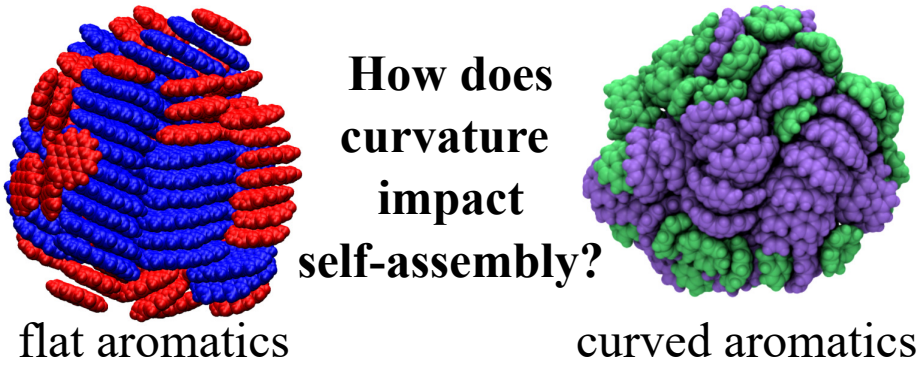
Chapter 8

Structure of clusters containing curved PAHs

The self-assembly and structure of nanoparticles containing curved polycyclic aromatic hydrocarbon molecules (cPAHs) are investigated using molecular modelling. These polar fullerene-like molecules are receiving increased attention recently due to the steric and electronic properties caused by the inclusion of five-membered ring(s) within their hexagonal lattice. Here the curPAHIP potential is extended to describe the interactions between large cPAHs. It is then applied to molecular dynamics simulations to produce equilibrated cPAH particles. Structural and energetic metrics, including diameter, density, intermolecular spacing, coordination number, alignment angle, radial distance, and energy values, are used to analyse systems containing cPAHs of different sizes and ratios, and containing flat PAHs or ions. Homogeneous cPAH particles are more tightly packed than their flat PAH counterparts, with large cPAHs displaying stacked columnar configurations absent in nanoparticles containing small cPAHs. Mixing cPAHs of different sizes disrupts the ordered mesophase and forms a core-shell structure in which the larger molecules make up the core and the smaller molecules comprise the shell, although this partitioning is less distinct compared to flat PAHs. In addition, the presence of flat PAHs and ions within cPAH nanoparticles promotes distinct arrangements dominated by weak dispersive interactions and strong electrostatic interactions, respectively. This provides insight into nanoparticles containing curved carbons relevant to applications in materials and combustion science.

Highlights:

- Self-assembly is primarily dependent on cPAH size rather than particle size
- Different cPAH sizes disrupt mesophase formation and arrange in a core-shell particle
- Systems containing cPAHs and fPAHs self-assemble into janus particles
- Cations promote or disrupt particle structure depending on the constituent cPAH size



The work presented in this chapter is being prepared for submission to Carbon.

8.1 Introduction

Curved carbon structures are found in many materials including porous carbons, glassy carbons, activated carbons [141, 218, 220], and combustion carbons [216]. For example, high resolution transmission electron microscopy of energy-relevant carbon materials such as coke and soot shows that a significant proportion (63% for young particles, 28-49% for mature carbons) of the constituent molecules are curved [350, 374, 216]. This curvature is predominantly caused by the presence of non-hexagonal rings, such as pentagons, within a hexagonal lattice [217]. The resulting molecules, known as curved polycyclic aromatic hydrocarbons (cPAHs), have steric and electronic properties not present in defect-free carbon materials containing hexagonal structures only (flat polycyclic aromatic hydrocarbons, fPAHs). In particular, the curvature redistributes electronic charge in the π-cloud and causes the molecules to possess a dipole moment due to the flexoelectric effect [221]. This allows curved molecules to interact in long-range electrostatic interactions not present in systems containing planar carbon molecules, while still retaining aromaticity and showing considerable electron delocalisation [130, 88].

These qualities cause the presence of curved aromatic molecules to influence material structure and properties. Curved molecules increase material porosity [141] and facilitate stronger adsorbate-adsorbent interactions [221] which, combined with high polarisability and high surface area, provide enhanced adsorption important for applications such as carbon sequestration, gas storage, and separation [286]. Curved aromatics also possess a combination of properties, including surface charge stabilisation, high charge mobility, significant dipole moment, and small band gap [223] that make them excellent candidates for applications such as optoelectronic devices, organic semiconductors, liquid crystals, electrodes, imaging probes, and batteries [272]. For example, integrating corannulene inside insulating porous scaffolds allows electronic properties to be tuned and results in a 10,000-fold conductivity enhancement [269]. Flame-formed carbon nanoparticles show quantum dot behaviour [206]; quantum dots have shown great promise in bioimaging as well as photovoltaic and light emitting applications due to their tunability, biocompatibility, luminosity, and solubility [371].

Accurately describing and characterising the self-assembly and nanostructure of curved carbon materials is of interest to processes producing these desirable materials and also more broadly since curvature is easily integrated even when unintended. Many ubiquitous materials, such as combustion-produced pollutants [216], interstellar medium [210], and graphite synthesised from mesophase pitch, are only fully characterised when the influence of curvature is understood. In particular, the degree of molecular alignment to form columnar

or stacked structures plays a significant role in the mechanical and electronic properties of materials, such as high electron transport characteristics desirable for organic electron devices [349] and the generation of graphisiting material. Previous work, including the results from Chapter 4, shows that fPAHs assemble in a nanocluster with good molecular alignment, forming an ordered mesophase [63]. In contrast, there is some evidence that the presence of curved molecules contributes to a low degree of molecular ordering in a material [374] and prevents graphitisation of carbonised material by disrupting the formation of the mesophase [5].

Previous work on the structure and properties of materials containing polycyclic aromatic hydrocarbons has focused on fPAHs [132, 64, 267, 152], with less attention given to cPAHs. Electronic structure calculations show that there are significant interactions between nested concave-to-convex homogeneous cPAH dimers [316, 49] and, as with fPAHs, cPAH interactions are dominated by π-π dispersion interactions with weaker contributions from CH-π electrostatic interactions. That being said, electrostatics are more significant for cPAH interactions compared to non-polar fPAHs due to their permanent dipole moments [49, 171]. Different degrees of curvature result in increased or decreased cPAH dimer strengths, depending on the interplay of geometry and electrostatic effects. Curvature is able to increase interaction strength by decreasing C-C distances for increased dispersion interactions [183] but very curved molecules can also experience increased steric hindrance [316, 217] and increased exchange-repulsion that serve to destabilise the dimer [183, 316].

X-ray crystallography and density functional theory calculations have shown that the crystal structure of cPAH systems are determined by an interplay of electrostatic and dispersive forces, but predicting the packed structure of cPAHs is not straightforward. A molecule's dipole moment and molecule bowl depth are identified as significant factors, but do not have clear threshold values that guarantee particular molecular arrangements [110]. In addition, the size [363, 111], curvature [256, 46, 282], rigidity [315, 349], functionalisation [285], and atomic composition [167] of cPAHs are known to influence their ability to form columnar stacks in solid state. These systems often show large π-π overlap and staggered stacked interactions to produce extended π networks enhanced by CH-π interactions.

Preliminary work on larger molecular systems, described in Chapter 7, suggests that the self-assembly of homogeneous cPAH clusters is significantly different from similarly sized fPAH clusters, which may be due to the ability of polar cPAHs to engage in electrostatic interactions. Computational studies show that the binding energies between fPAHs of different sizes are weaker than those within a homogeneous system containing one molecule size [267]. This heterogeneity decreases the stability of a nanoparticle containing different molecule sizes

and leads to a distinct partitioning in which the larger molecules formed the cluster core and the smaller molecules resided in the outer shell, shown in Chapter 4. The extent to which this nanostructure is also seen within cPAH systems has not yet been investigated, but dimer calculations suggest that bowl complementarity may produce different behaviour by enhancing the stability of heterogeneous cPAHs clusters [49]. Previous work has also characterised the first solvation shells of fPAHs around an alkali-metal ion [21, 65], but this has not been well-explored for cPAHs, which would likely self-assemble differently due to their polarity. It is therefore of great interest to understand how the fundamental interaction differences in shape and binding behaviour between fPAHs and cPAHs may influence their self-assembly in homogeneous and heterogeneous nanoparticles. To date, detailed studies of cPAHs have primarily included electronic structure calculations or crystal structure experiments, as described above, neither of which provide information about intermolecular dynamics and particle nanostructure. Previous work looking at systems containing cPAHs and fPAHs show the effect of structure on porosity and adsorption [370, 80], but do not evaluate the system dynamics that are crucial to understand self-assembly and structural properties. In addition, most work does not include the flexoelectric effects so long-range dipole-dipole electrostatic interactions are not included.

The purpose of this chapter is to explore the self-assembly and internal nanostructure of clusters containing cPAHs, with the aim of answering the following questions relevant to carbon scientists: *What is the energy and structure of cPAH nanoparticle systems and how do these differ from systems containing fPAHs? What is the influence of particle size, molecule size and proportion, and presence of ions or fPAHs?* These are addressed by extending a force field parameterised for cPAH systems and using it within molecular dynamics simulations to provide a detailed assessment of cPAH cluster self-assembly. This analysis provides insight into the molecule interactions and nanoparticle structure that are important for better understanding carbon nanomaterials relevant to applications such as batteries, adsorbents and optoelectronics.

8.2 Methods

8.2.1 Systems

To thoroughly evaluate the self-assembly of nanoparticles containing cPAHs, three different molecule types with different sizes and degrees of curvature are considered: a small fPAH containing seven hexagonal rings (coronene, $C_{24}H_{12}$), the smallest cPAH which contains

one central pentagonal ring with five surrounding hexagonal rings (corannulene, $C_{20}H_{10}$), and a larger cPAH containing two embedded pentagonal rings based on HRTEM analysis of early soot nanoparticles [216] ($C_{42}H_{14}$), all shown in Figure 8.3. The notation X_y is used to describe the clusters studied, where X values refer to the molecule types and y gives the number of each type of molecule within the cluster. The molecule species considered in this chapter are: corannulene (indicated as A), $C_{42}H_{14}$ (B), coronene (C), circumcoronene (fPAH $C_{54}H_{18}$) (D), and potassium cation (K). For example, $A_{50}B_{50}$ indicates a cluster containing 100 molecules, made up of 50 corannulene and 50 $C_{42}H_{14}$ molecules.

The influence of particle size, molecule size, molecule curvature, molecule ratio and ion interactions are evaluated by considering 17 different clusters. Homogeneous clusters containing 25, 50, 100 and 200 of A and clusters containing 25, 50 and 100 of B are studied to provide information across molecule and cluster sizes. A series of heterogeneous clusters each containing 40 molecules of different sizes and proportions (A and B in proportions of 30:10, 20:20 and 10:30 are studied) allows evaluation of heterogeneity effects. In addition, a heterogeneous cluster containing 50 A and 50 B addresses heterogeneous cluster size effects. Clusters containing 20 A and 20 C provide insight into the interactions between cPAHs and fPAHs, and clusters containing 40 A or B with one or two potassium cations allow investigation into the self-assembly of ion-containing clusters. Snapshot images of each cluster considered in this chapter are shown in Figure 8.3. It should be noted that the terms (nano)particle and cluster are effectively synonymous: a nanoparticle is a cluster of molecules.

8.2.2 Force field development

Before any simulations can be conducted, the suitability of an intermolecular potential must be thoroughly evaluated. The isoPAHAP potential is an all-atom isotropic intermolecular description developed for fPAHs [324]. It was developed using high accuracy benchmark quantum calculations and has been used in dynamic and stochastic simulations of fPAH systems [327, 132, 252]. This potential uses fixed atom-centred charges, which are suitable for fPAHs where electrostatic interactions arise principally from the terminal groups. It is not able to capture local dipole moments located at strained internal carbon sites within cPAHs. Therefore, as outlined in Chapter 7, a new atomic potential for cPAHs was developed, called the curPAHIP potential. The curPAHIP potential models the increased polarity of cPAHs using a modified molecule description with off-site point charges located above the pentagonal carbon atoms and optimised potential parameters parameterised to SAPT(DFT) energies.

Previous work introducing the curPAHIP potential included the cPAH A only. Here the molecular description is extended to include the larger cPAH B and assess the suitability of the curPAHIP potential for the new systems. Following the parameterisation method developed for A in Chapter 7, the geometry of B is minimised and massless charges are added 0.052 nm above each of the pentagonal carbon atoms to match the calculated dipole moment of 5.28 D [216]. The resulting atomic coordinates and charges of the minimised B monomer are provided in Appendix C.

The binding energies of cPAH dimers are calculated to assess the suitability of curPAHIP in describing systems containing larger cPAHs. Density functional theory (DFT) calculations performed with the Gaussian 16 software [121] are used to determine the geometry of cPAH monomers and geometries and energies of cPAH dimers in order to assess the atomic and intermolecular descriptions employed here. The B97-D functional [133], a hybrid functional with an empirical dispersion correction, is used following a methodology similar to that used by Martin et al. [217]. Geometry optimisations are carried at the B97-D/6-311G(d,p) level of theory with frequency calculations to ensure suitable minima are found, and subsequent single point calculations use B97-D/cc-pVTZ, with basis set superposition error corrections. These functional and basis sets were selected since they provide good agreement (within 4%) with benchmark CCSD(T) [171] and SAPT(DFT) [49] calculations of small cPAH molecules.

Binding energies are shown for four dimers in Figure 8.1, compared with intermolecular potential curves calculated using the software ORIENT [311]. The energy computed with the isoPAHAP potential is included to highlight the behaviour of a potential, developed for fPAHs, that does not include the enhanced electrostatics and dispersion due to the flexoelectric effect and increased polarisability, respectively, and in all cases the isoPAHAP potential significantly underestimates the binding energy and overestimates the dimer distance at the minimum energy. The curPAHIP potential agrees well with the DFT and SAPT(DFT) results for the A dimer, see Figure 8.1(a). This is expected since these *ab initio* values were used in the parameterisation of the curPAHIP potential. Figure 8.1(b) shows that the curPAHIP potential can be extended to cPAH molecules larger than A, since there is good agreement (within 5% of the dispersion-corrected DFT energies) for the larger B. In contrast, the isoPAHAP gives a minimum energy value that is 31% smaller than the DFT calculations. The energies of heterogeneous dimers, containing one A and one B, are also well captured by the curPAHIP potential, as seen in Figure 8.1(c) and (d). The repulsive branch of the curPAHIP potential is slightly shifted to smaller distances than the DFT values in some cases, which is acceptable since the repulsive branch still possesses the same mathematical form that has been found to produce appropriate cluster morphologies [252].

An understanding of the potential error in structural metrics calculated from molecule arrangements caused by the use of the curPAHIP potential can be obtained by comparing to the DFT results. As the molecules used in the parameterisation of the potential, the A dimer shows no significant deviation. For dimers containing B, there is a slight overall shift of around 0.02 nm between the curPAHIP potential energies and those calculated by DFT, which provides an estimate of the error in the intermolecular spacings determined in simulations using the cPAH intermolecular potential.

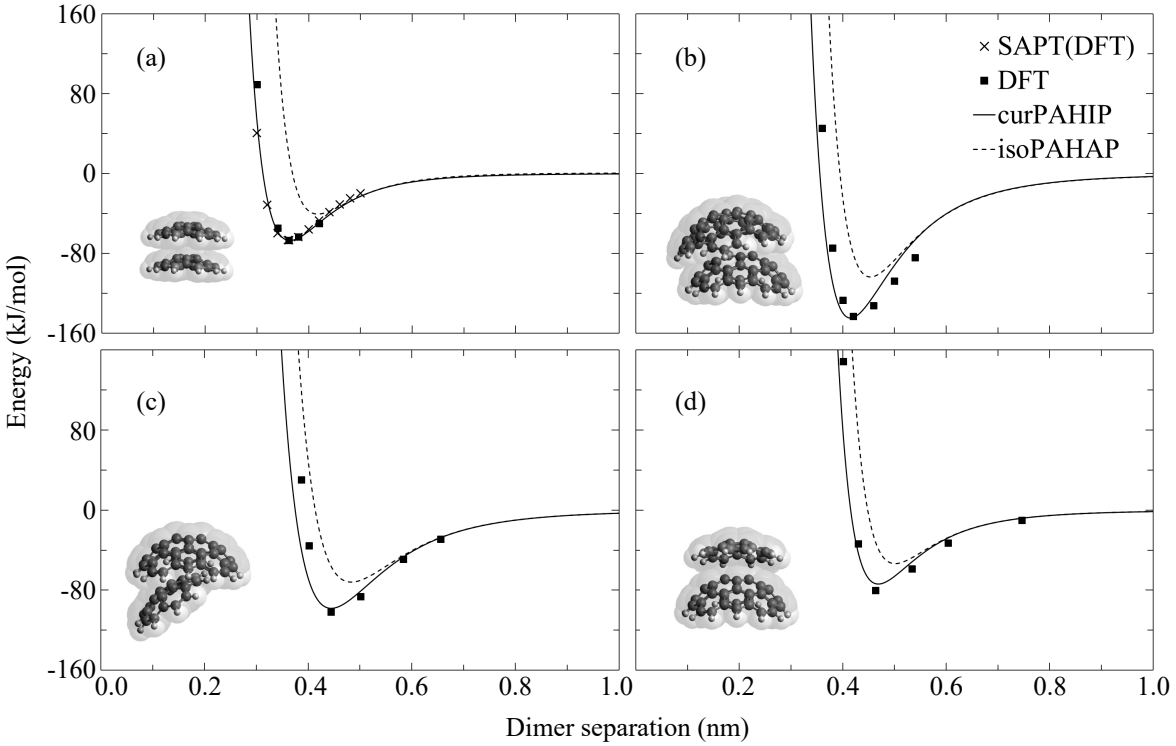


Fig. 8.1 Interaction energy versus separation distance for cPAH dimers determined from SAPT(DFT) calculations [49], DFT calculations, the curPAHIP potential and the isoPAHAP potential. The dimers are as follows: (a) two A, (b) two B, (c) and (d) one A and one B each.

Figure 8.1 focuses on sandwich type interactions between cPAH dimers although the nested dimer with the smaller A on the concave side of the larger B (Figure 8.1(c)) shows some T-shaped interaction. The ability of these intermolecular potentials to capture CH-π interactions is further explored with the T-shaped dimers of A and C in Figure 8.2. These interactions are weaker than their sandwich counterparts but there is good agreement between the potentials and the electronic structure calculations. The A dimer in which the molecules interact at the convex surface (Figure 8.2(a)) shows the same behaviour as the C T-shaped dimer. In addition, the isoPAHAP and curPAHIP potentials show similar trends (within 7% of each other) for all T-shaped dimers. This suggests that the increased interaction strengths and decreased

binding distances observed for the dimer in which the molecules interact at the concave surface (Figure 8.2(b)) compared to the dimer involving interactions at the convex surface (Figure 8.2(a)) are due to sterically-enhanced dispersion rather than enhanced electrostatic interactions from the flexoelectric dipole. Importantly, the T-shaped A dimer at the concave surface is stronger by 10 kJ/mol (over 30% enhancement) than the T-shaped interaction of fPAH C. This enhanced interaction allows T-shaped interactions to be energetically favoured over sandwich interactions when many corannulene molecules are considered, such as in a cluster or bulk system.

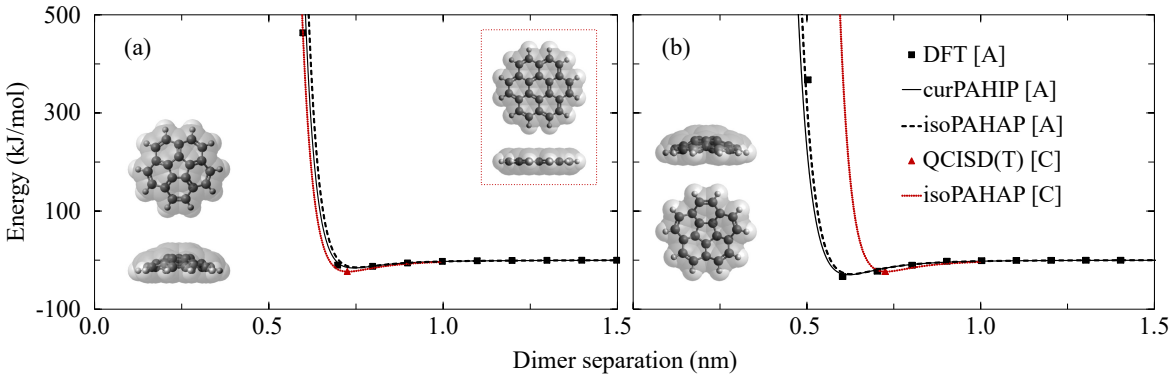


Fig. 8.2 Interaction energy versus separation distance for cPAH and fPAH dimers in T-shaped configurations determined from QCISD(T) calculations [170], DFT calculations, the curPAHIP potential and the isoPAHAP potential. Square brackets indicate the corresponding molecule types for each calculation method. The T-shaped dimer interactions with (a) the convex surface and (b) the concave surface of A are compared with the T-shaped C dimer, which is pictured in the dotted inset.

Simulations considering both fPAHs and cPAHs in a single cluster used the curPAHIP potential to describe the mixed molecule interactions and simulations. Using the isoPAHAP potential to describe these interactions showed very similar results.

8.2.3 Molecular dynamics

Cluster configurations are produced using a multi-step molecular dynamics simulation process. Clusters are initialised in a mixed configuration, with molecules randomly placed within a large spherical volume using the PACKMOL software [222]. Excess energy is removed by an energy minimisation step using the steepest descent algorithm, followed by the low-memory Broyden-Fletcher-Goldfarb-Shanno method [48].

Replica Exchange Molecular Dynamics (REMD) simulations are used to rapidly produce equilibrated cluster systems. Modelled after Monte Carlo parallel tempering [161], REMD is

an advanced form of molecular dynamics that involves evaluating simultaneous isothermal systems, called replicas, across a temperature range [313]. At regular intervals throughout the simulation, neighbouring replicas are able to exchange spatial information based on a Boltzmann-weighted temperature dependent probability. This method allows efficient and rapid sampling of the potential energy landscape of the system since low energy replicas are able to explore new configurations generated by higher energy replicas.

REMD simulations are conducted for 3 ns. A large temperature range is selected to include solid-like and liquid-like particle morphologies. For clusters containing only A, only B, A and B, and A and C this corresponds to temperature ranges of 200–800 K, 400–1600 K, 200–1600 K and 200–800 K, respectively. These require between 27 and 84 REMD replicas to maintain an acceptable replica exchange acceptance. As in previous work, replica temperatures for REMD simulations are selected using an exponential temperature distribution given in Equation 4.1. The j and m parameters, given in Table 8.1, are selected for each cluster system to allow satisfactory exchange between replicas using these temperature ranges. Exchange acceptances for simulations conducted in this chapter have an average between 0.20 and 0.35 with small system ranges, which indicates robust replica exchange with a good balance between equilibration within replicas and exchange between replicas. Replica exchange attempts are made every 100 fs, since frequent exchanges have been shown to increase efficiency without affecting the ensemble being sampled [301, 302].

As described in detail in Chapter 4, a flat-bottomed spherical position potential is applied within the REMD simulations to address complete evaporation of small molecules from the cluster at high temperatures. Individual 1 ns simulations using classical molecular dynamics (MD) are then conducted at each desired temperature from the final REMD configurations. No position potential is implemented for these post-REMD MD simulations.

In the REMD and MD simulations the NVT ensemble, where a constant number of atoms, system volume and temperature are maintained, is sampled using a chain of 10 Nosé-Hoover thermostats for temperature control. A velocity Verlet integrator [336] is used to advance the configuration in 1 fs time steps and all simulations are conducted *in vacuo* without periodic boundary conditions. Intramolecular forces are determined using the OPLS-AA force field [179] for molecular bonds, angles, dihedral and improper dihedral angles. The curPAHIP intermolecular potential developed in Chapter 7 is used to describe interactions between cPAHs and intermolecular cut-offs are set to 3.0 nm. All minimisation, REMD and MD simulations are conducted using GROMACS 5.1.4 [4]. Purpose-made scripts are used to the process the output and the software VMD [162] provides visualisations.

Table 8.1 Replica temperature selection parameters and temperature ranges (in K) for cPAH clusters considered in this chapter.

Cluster	m	j	lowest T	highest T
A_{25}	0.054	0-26	200	≥ 800
A_{40}	0.050	0-29	200	≥ 800
A_{50}	0.040	0-35	200	≥ 800
A_{100}	0.025	0-59	200	≥ 800
A_{200}	0.017	0-86	200	≥ 800
B_{25}	0.045	0-32	400	≥ 1600
B_{40}	0.032	0-44	400	≥ 1600
B_{50}	0.030	0-47	400	≥ 1600
B_{100}	0.024	0-59	400	≥ 1600
$A_{20}B_{20}$	0.030	0-71	200	≥ 1600
$A_{50}B_{50}$	0.025	0-83	200	≥ 1600
$A_{30}B_{10}$	0.036	0-59	200	≥ 1600
$A_{10}B_{30}$	0.028	0-74	200	≥ 1600
$B_{40}K_1$	0.030	0-47	400	≥ 1600
$A_{40}K_1$	0.050	0-29	200	≥ 800
$A_{40}K_2$	0.050	0-29	200	≥ 1600

8.2.4 Structural analyses

A number of different metrics, including density, intermolecular spacing, coordination number, alignment angle and radial distance, are used to evaluate system structural properties. All metrics are averaged over the final 500 ps of the post-REMD MD simulation of the lowest temperature replica (*i.e.* 3.5–4.0 ns of total simulation time) using a timestep of 1 ps. Many calculations require the identification of near neighbours for each molecule within the system. For this, molecules are considered neighbouring if their centres of mass are within the cut-off radius R for at least half of the 500 ps production period. Unless otherwise stated, values of R are selected to allow $> 85\%$ of molecules to have at least one identified neighbour. This results in cut-off radii of $R_A = 0.7$ nm, $R_B = 0.5$ nm, and $R_C = 0.5$ nm for A, B and C, respectively. The sensitivity of the cut-off values on calculated cluster properties is discussed in Appendix C.

Diameter and density values are determined using the solvent-excluded surface of the cluster calculated using a rolling sphere algorithm [284], as in previous work of fPAH clusters [62] (given in Chapter 5). This provides values that are more accurate than commonly used spherical approximations since they are directly calculated from the three-dimensional surfaces. Molecular alignment angles are calculated to provide information on the relative

configurations of neighbouring molecules within the clusters studied. An alignment angle is defined as the angle between normal vectors to the central rings (for A this is the pentagonal ring and for B this is the central hexagonal ring) of the neighbouring cPAHs considered. Appendix C provides a schematic of the alignment angle between two neighbouring A molecules. A quantitative measure of the degree of stacking order in the molecular structures is provided through the use of coordination numbers (CNs), calculated as the number of near neighbours averaged over each molecule type (see Equation 6.2). To consider only π-π stacking interactions within this metric, values of R are selected to include sandwich-type stacked interactions between molecules but exclude molecules more than one layer away. Dimer distances at the minimum energy provide the minimum R values for each molecule type as $R_{\mathrm{A}} = 0.4$ nm and $R_{\mathrm{B}} = 0.5$ nm. Radial distance is defined as the distance between a molecule type and the cluster centre averaged over all atoms (see Equation 6.1) and provides insight into the spatial partitioning of molecule types within a cluster. This metric is particularly useful for understanding the self-assembly of heterogeneous systems.

8.3 Results and Discussion

Figure 8.3 shows all of the low energy cluster geometries and Table 8.2 provides a summary table containing the corresponding structural metrics. The discussion of these results will be structured around questions in materials and combustion science including:

- **Do cPAHs self-assemble into an ordered phase?** cPAHs clusters are analysed to explore the impact curvature has on the development of an ordered mesophase.
- **What is the internal nanostructure of cPAH nanoparticles?** cPAH clusters are evaluated with additional metrics (densities, radial distances and energetics) to explore properties particularly relevant to combustion-generated nanoparticle pollutants and nanoparticle synthesis.
- **How do complex cPAH systems self-assemble?** Clusters containing cPAHs with fPAHs or ions are structurally and energetically analysed to provide insights into real-world systems such as janus nanoparticles, battery materials and combustion-generated nanoparticles.

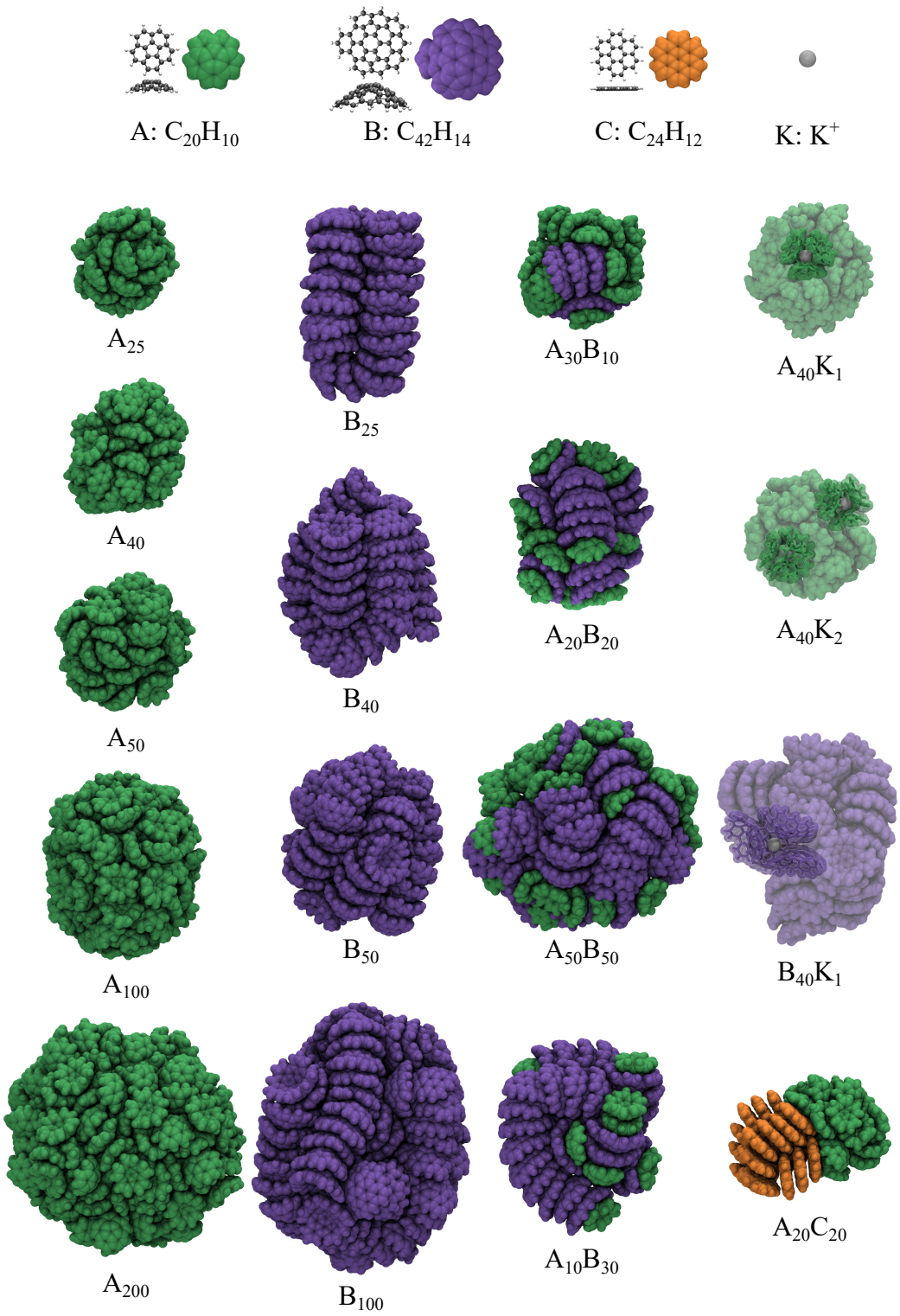


Fig. 8.3 Visualisations of the corannulene molecule (A, coloured green), $C_{42}H_{14}$ molecule (B, coloured purple), coronene molecule (C, coloured orange) and potassium cation (K, coloured grey), and clusters studied in this chapter. Ion-containing clusters are shaded to emphasise the solvation shell surrounding the ion(s).

Table 8.2 Cluster diameter (nm), density (g/cm^3), intermolecular energy (kJ/mol per molecule), intermolecular spacing (nm), coordination number, and radial distance, r, of molecule A and molecule B (nm) for all clusters studied in this chapter.

Cluster	Diameter	Density	Energy	Spacing	CN	r_A	r_B
A_{25}	2.36	1.50	-75	0.59	0.01	0.91	–
A_{40}	2.77	1.49	-82	0.59	0.02	1.07	–
A_{50}	3.00	1.47	-84	0.60	0.01	1.14	–
A_{100}	3.79	1.45	-92	0.59	0.00	1.46	–
A_{200}	4.79	1.45	-97	0.60	0.01	1.85	–
B_{25}	2.96	1.59	-144	0.44	1.57	–	1.27
B_{40}	3.49	1.55	-152	0.45	1.58	–	1.45
B_{50}	3.76	1.55	-157	0.45	1.43	–	1.49
B_{100}	4.77	1.52	-165	0.45	1.36	–	1.92
$A_{10}B_{30}$	3.31	1.58	-147	0.44	1.04	1.53	1.30
$A_{20}B_{20}$	3.16	1.55	-125	0.44	0.92	1.31	1.20
$A_{50}B_{50}$	4.31	1.52	-135	0.51	0.70	1.85	1.49
$A_{30}B_{10}$	2.98	1.52	-103	0.54	0.74	1.21	1.04
$A_{20}C_{20}$	2.88	1.46	-80	0.43	0.74	1.18	1.21 (r_C)
$A_{40}K_1$	2.78	1.49	-88	0.61	0.10	1.06	–
$A_{40}K_2$	2.79	1.48	-94	0.58	0.20	1.05	–
$B_{40}K_1$	3.50	1.54	-149	0.46	1.21	–	1.35

8.3.1 How do cPAHs self-assemble?

As mentioned, mesophase formation (the molecular alignment of aromatic molecules) is critical for graphitisation. The intermolecular spacing, coordination number and alignment angle values of molecules within homogeneous and heterogeneous clusters of cPAHs will be compared with similar clusters of fPAHs and experimental systems to provide insights into this question regarding cPAH self-assembly.

Average intermolecular spacing is an important experimental quantity when tracking the formation of a mesophase and subsequent graphitisation as well as the presence of curvature in combustion carbons [41]. The homogeneous cPAH clusters evaluated have intermolecular spacings that do not change significantly with cluster size but depend strongly on the constituent molecule size. Clusters containing A possess an average intermolecular spacing of 0.59 nm while clusters containing B have an average spacing of 0.45 nm. This shows that the spacings within homogeneous clusters are controlled by the molecular composition of the cluster rather than the number of molecules in a cluster. Intuitively, one would expect the intermolecular spacing to increase with molecule curvature due to steric effects that

prevent the molecules from interacting closely, however the opposite trend is observed. This suggests that these two cPAH molecule types configure in different arrangements that are not controlled solely by steric effects.

To provide a comparison to reference molecule arrangements, the spacings of cPAH dimers using electronic structure calculations from Figure 8.1 are (a) 0.36 nm, (b) 0.42 nm, (c), 0.44 nm, and (d) 0.46 nm. This shows that the average spacing of A within a homogeneous cluster is significantly larger than that of its sandwich dimer. The average intermolecular spacing within a single layer of the A crystal structure characterised by X-ray crystallography is 0.57 nm, suggesting the cluster structure of A likely possesses significant contributions from CH-π interactions, in which the more positive region at the rim of one molecule is almost perpendicular to the negative region and the bottom of another molecule in a T-shaped configuration, rather than tight sandwich interactions. Further information on the A crystal structure is provided in Appendix C. In contrast, the average intermolecular spacing of a B cluster is similar to that of its sandwich dimer and similar to the spacing found in crystal structures of similarly curved indenocorannulene species (0.34–0.37 nm [110]). The intermolecular spacings within clusters containing B do not change readily with the cut-off distance used, indicating that B have distinct near neighbours, for example in a highly stacked configuration. In contrast, A spacings are strongly correlated to the cut-off distance selected, suggesting that the molecules are not arranged in structured stacked layers. Intermolecular spacings are reported for a number of cut-off distances in Appendix C.

The intermolecular spacings within heterogeneous cPAH clusters suggest that both the molecular ratio and cluster size play a role in the average spacing. This is investigated further in Table 8.3, which breaks down the average results by considering the molecule type contributions to the intermolecular spacing. The interactions between molecules of the same type maintain spacings similar to those in the homogeneous clusters, so that the spacings between A are larger than those between B. As in the homogeneous clusters, the B-B spacings are similar to that of a minimised B dimer and reasonably uniform across all heterogeneous clusters. Cluster size does not influence the A-A spacing, however the proportion of A plays a role. The mixed molecule interactions have similar spacings to those of the larger cPAH suggesting that B promotes the close stacking behaviour of π-π interactions with A. This is more pronounced in clusters with a higher proportion of B compared to A and is also cluster-size dependent. The differences between average spacings for the heterogeneous clusters (for example comparing $A_{20}B_{20}$ and $A_{50}B_{50}$) are therefore observed to be largely due to the number of neighbouring A or B within each cluster rather than the molecule type intermolecular spacings. This molecule type behaviour is likely due to the difference in curvature between these two cPAHs rather than the molecule sizes alone, since fPAH clusters

containing a disparity of molecule sizes do not possess these differences (for example, a cluster containing 16 C and 16 D has an average spacing of 0.42 nm, with C-C spacings of 0.43 nm, C-D of 0.40 nm, and D-D of 0.41 nm),

Table 8.3 Intermolecular spacings within heterogeneous clusters, considering average distances between molecule types A and B.

Cluster	B-B	B-A	A-A
$A_{10}B_{30}$	0.43	0.43	0.53
$A_{20}B_{20}$	0.43	0.41	0.61
$A_{30}B_{10}$	0.44	0.48	0.58
$A_{50}B_{50}$	0.44	0.47	0.60

These intermolecular spacings are significantly increased compared to planar carbons, especially for heterogeneous clusters. For example, a cluster containing fPAH 100 C [63] has an average intermolecular spacing of 0.43 nm while all heterogeneous clusters have spacings $\geq$0.44 nm. In flame-produced soot particles similar spacings are found (0.38–0.48 nm depending on the particle maturity using high resolution transmission electron microscopy [41, 15]) to those seen for B and heterogeneous cPAH clusters. These results suggest that A or homogeneous clusters are not complete representations of these experimental systems and the intermolecular spacing is a useful quantity for detecting the presence of curvature in nanoparticles and carbon materials.

Coordination number values provide information on the extent of stacking interactions, which can help explain the intermolecular spacing results by identifying near neighbour patterns. In particular, this metric shows whether T-shaped interactions (in which molecules do not possess any near neighbours) or sandwich interactions (in which each molecule has one or two near neighbours) dominate. The homogeneous A clusters show an average CN of 0.01 ± 0.01, indicating predominantly T-shaped interactions, while the B clusters possess a CN of 1.48 ± 0.09, suggesting that they self-assemble with sandwich interactions. As seen in the cluster snapshots, B interact closely such that each molecular bowl inserts into the concave surface of its neighbour, allowing each molecule to possess on average more than one near neighbour in a stacked configuration. This is very similar to the arrangements of fPAH within clusters (for example, C_{100} has an average CN of 1.6 [64]) and cPAH hybrids that form tight stacks [97]. In contrast, A within clusters do not have near neighbours within stacking distance and the molecular bowls do not pack tightly within each other. As discussed by Liu et al. [206], Adkins et al. [6], molecular stacking is an important factor for the band gap of a PAH nanoparticle.

To further examine the influence of compositional heterogeneity, CNs are compared for different clusters each containing 40 cPAH molecules in Figure 8.4. The molecule-specific CNs within these heterogeneous clusters align with those calculated for homogeneous clusters. The B within all clusters have CNs above 1, indicating that on average these molecules have more than one near neighbour in a stacked arrangement. In contrast, all A have CNs significantly below 1, which shows that these molecules do not arrange in close stacks. This suggests that the formation of an ordered phase is more likely with larger cPAHs compared to small species such as A. This molecule type difference, where smaller molecules possess lower CNs than larger molecules within a cluster, is also seen in fPAHs (for example, CNs of 2.00 and 1.25 for the C and D, respectively, within a $C_{16}D_{16}$ cluster) and is linked to the presence of smaller PAHs on the cluster surface (often by capping the ends of molecule stacks) compared to the bulk-residing larger PAHs. The CNs for both cPAH sizes follow the trend $A_{40} < A_{30}B_{10} < A_{20}B_{20} < A_{10}B_{30}$ (shown with a solid arrow), illustrating that the addition of B into A clusters increases the degree of order and stacking.

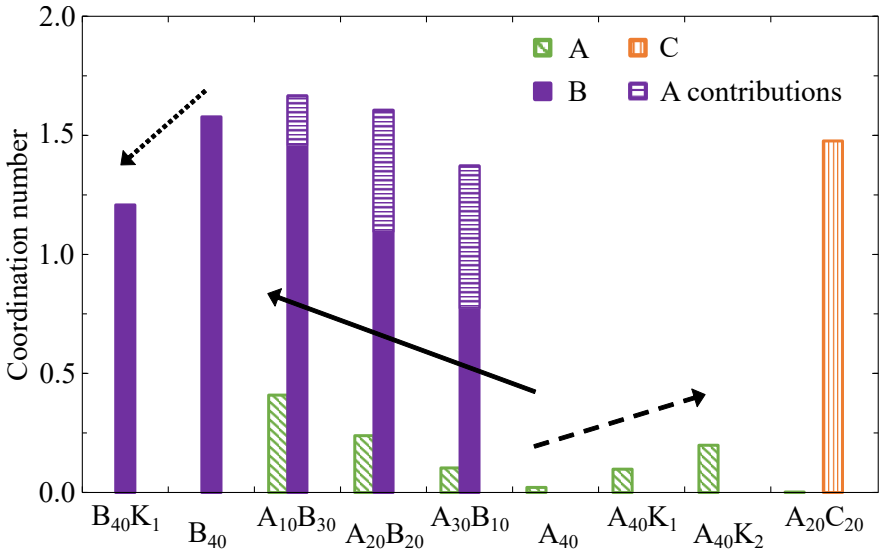


Fig. 8.4 Molecule type coordination number values for clusters containing 40 PAHs.

Understanding which molecule types are present as near neighbours within these heterogeneous systems is also of interest. It is observed that the CNs of A have no contribution from A, meaning that all A stacking interactions come from neighbouring B. The majority of B stacking interactions come from other B, however some of the interactions are with A in proportion to the molecular ratios within the cluster, shown as insets with horizontal lines in Figure 8.4. This highlights that the smaller cPAH A interacts most strongly in the concave surface of the larger cPAH B, in agreement with calculated DFT dimer energies and previous suggestion that interactions between bowl-complementing heterogeneous cPAH dimers

could remove the strict steric constraints limiting the interaction strengths of homogeneous cPAHs [217].

Figure 8.5 presents histograms of the molecular alignment angles, illustrating the relative configurations of neighbouring molecules within the homogeneous clusters studied. It is again immediately clear that the two molecule sizes behave differently but these trends are consistent across all cluster sizes considered (additional cluster size plots can be seen in Figures 8.6, C.1 and 8.7). The systems containing B (in the top row) show a single significant peak around 20°. This suggests that nearly all of the molecules interact in a tilted stack formation, forming one dimensional columns in which the molecule bowls possess a concave-to-convex alignment with dipole moment vectors nearly aligned. This small alignment angle provides significant π-π overlap that, within columnar structures, is identified as a feature of materials with good performance in organic electronic devices. This dominant alignment angle is the same as that observed between a minimised B dimer (shown as a dashed vertical line), indicating that this is a stable arrangement. Similar self-assembled structures are observed in crystals containing indenocorannulene molecules [110] and other highly curved cPAHs that form polar crystals with strong photoluminescence [66]. Neighbouring columnar stacks of B have opposite bowl directions and weak CH-π interactions provide little interaction between columns. Interestingly, the columns show significant tilting at their ends such that parallel columns curve to be nearly continuous, which is not a feature of crystal structures and likely influences cluster surface properties. fPAH clusters, such as the representative C_{100} shown in Figure 8.5, show even sharper angle peaks, suggesting a strong mesophase formation. However, these peaks are centred around 0° and 180°, which both correspond to aligned molecular planes (since the alignment angle calculation uses a normal vector). The variance in the fPAH alignment angles are likely from shifted interacting molecular planes as well as small tilting angles. This π-π stacking arrangement combines with CH-π interactions to produce a herringbone-like structure within both clusters and bulk crystal fPAHs [184], without the column curving present in cPAH clusters.

In contrast, clusters containing A (bottom row of Figure 8.5) show a dominant but broad peak around 45°, with smaller wide peaks around 130° and 170°. This agrees with previous work that shows A do not pack with any long-range order [139, 255, 180, 349, 294] (the angles found in the crystal structure are shown as vertical dashed lines), with some CH-π interactions but limited π-π interactions. As seen in the dimer energies, the molecule shapes enhance the interaction energy of T-shaped dimers in these small cPAHs so that, in contrast to the preferred stacked structure of fPAHs, this configuration becomes favoured for extended systems. It is interesting that this bulk molecular arrangement is observed even within the small nanocluster systems examined here. The method constraints of this work highlight that

this structure is not due to the rapid bowl inversion dynamics or induced polarities within the system but instead to the molecule size, shape and electrostatic properties.

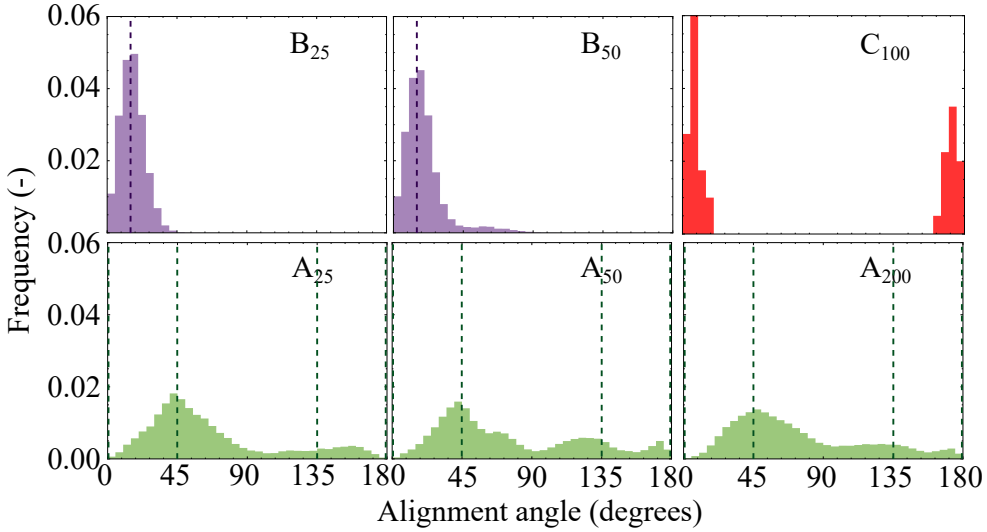


Fig. 8.5 Alignment angle distributions for homogeneous B and A clusters across different cluster sizes, with a fPAH cluster containing C included for comparison.

Given that the two molecule types self-assemble with different alignment angles, how mixing these molecules influences the alignment angle distributions is explored. Figure 8.6 (second row) shows heterogeneous clusters each containing 40 molecules alongside the corresponding homogeneous clusters (third row). Additional results from larger clusters are shown in Figure 8.7. In all heterogeneous clusters, the individual molecule alignment peaks align reasonably well with those seen in the homogeneous clusters, A crystal structure and B dimer, with the majority of B around 20° and many A around 45°. There is an additional small B peak around 45°, indicating some deviations from columnar stacking that align with the dominant arrangement of A. This is more pronounced in clusters containing $\geq$50% A, suggesting that A molecules are located within B columns. The molecule proportions also influence A angles, with shifts towards lower angles and reduced values around 130°. The $A_{30}B_{10}$ cluster shows a distribution similar to that of the homogeneous A cluster, although the peaks are much broader. As the proportion of B increases their influence is felt on A so that in the cluster containing equal ratios of both molecules ($A_{20}B_{20}$), the A peak at 45° is spread out towards 0° and the former 135° peak is pushed towards 170°. In the cluster containing 75% B ($A_{10}B_{30}$), the largest angle peak for A aligns with that of B at 20°. This indicates that A are stacking in a similar fashion to B, which can be seen in the corresponding cluster snapshot (Figure 8.3) where A predominantly interact individually at the ends of B stacks, fitting in the concave or convex surfaces. Although similar molecule size interactions are observed in heterogeneous fPAH clusters, these planar aromatics do not self-assemble

with different alignments depending on molecule size and thus heterogeneous clusters show the same highly aligned columnar stacking behaviour as homogeneous clusters, for example shown in the $C_{16}D_{16}$ cluster in Figure 8.6 (top row).

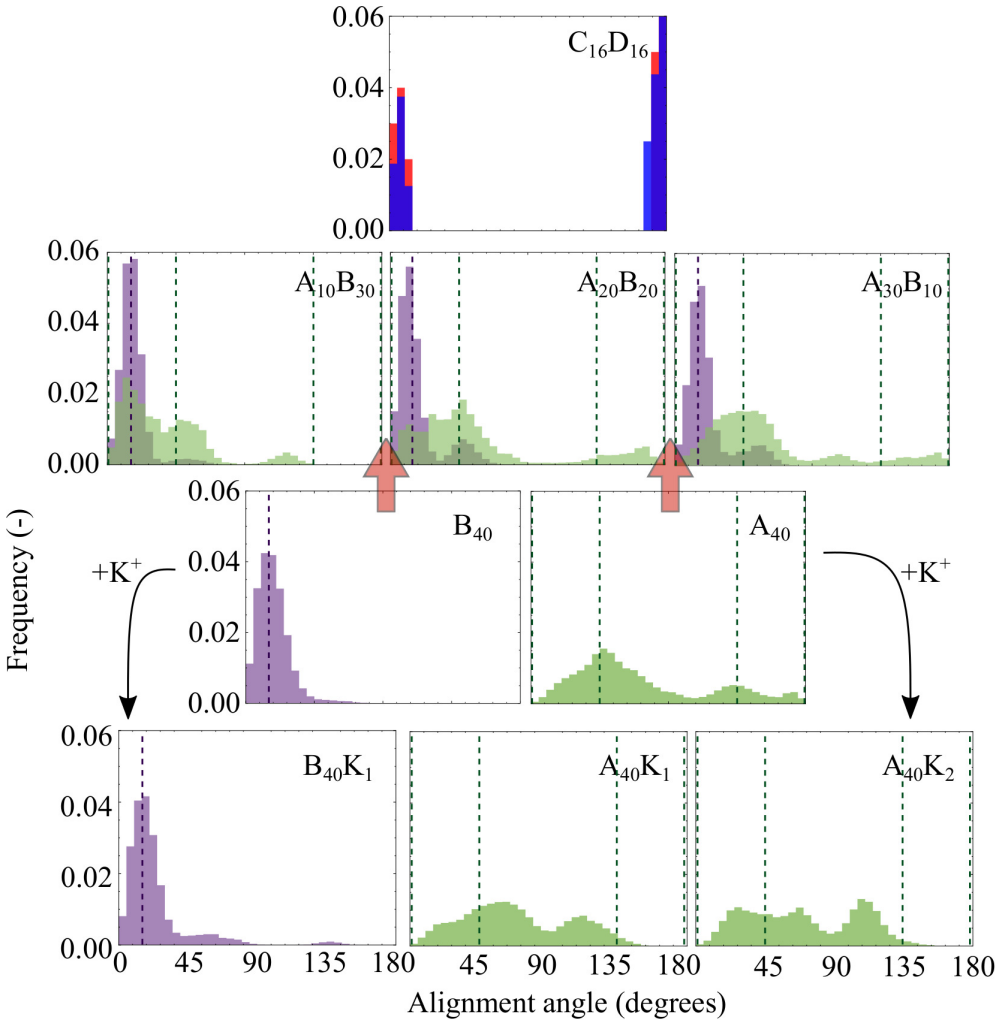


Fig. 8.6 Alignment angle distributions for A and B clusters containing 40 molecules. The third row shows the homogeneous clusters, with red arrows showing the heterogeneous clusters (second row) and curved arrows indicating ion-containing clusters (bottom row). Angle distributions for a heterogeneous fPAH cluster (top row) is provided for comparison.

This detailed evaluation of intermolecular spacing, CNs and alignment angles shows that the self-assembly of cPAHs depends strongly on their molecular geometry and not the cluster size. Strong CH-π attractions between molecule rims and bowl centres control the arrangement of A while dipole-dipole attractions and dispersion interactions control the nanostructure of B clusters. This molecule size dependent structure is observed within the crystal structures of cPAHs and also influences the behaviour of heterogeneous systems. Systems containing B

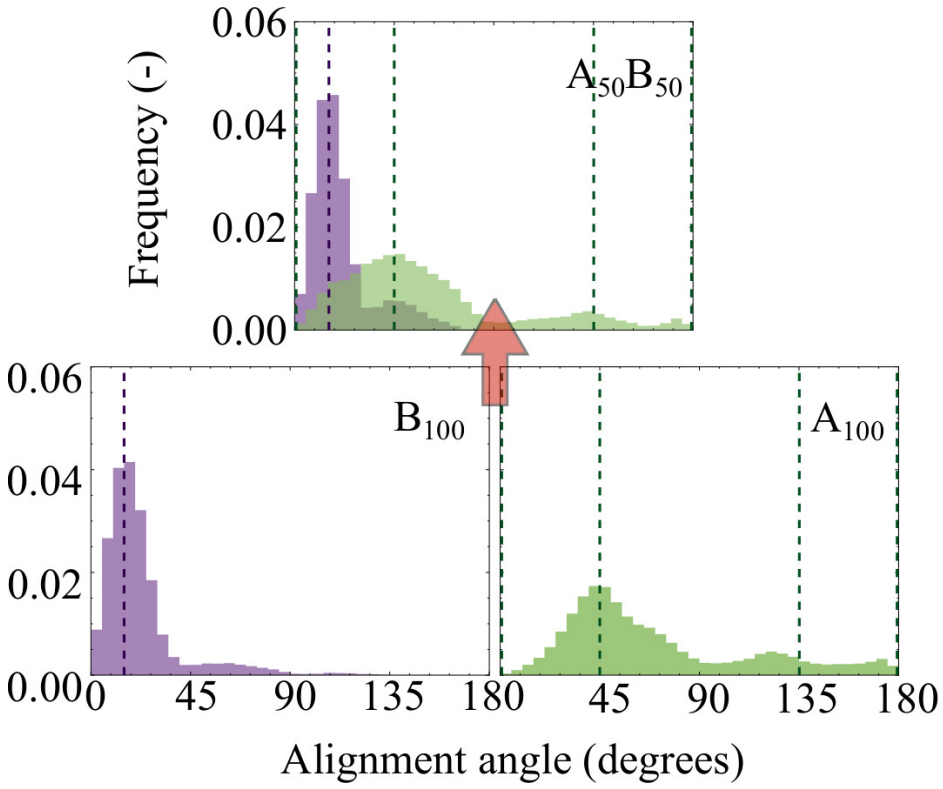


Fig. 8.7 Alignment angle distributions for homogeneous and heterogeneous A and B clusters each containing 100 molecules.

show high CNs and low alignment angles, suggesting good molecular alignment similar to clusters of fPAHs. However, cPAH clusters containing even small proportions of A disrupt mesophase formation.

8.3.2 What is the internal structure of cPAH nanoparticles?

Characterising the internal nanostructure of cPAH clusters provides valuable information relevant to the formation and composition of combustion-generated nanoparticles as well as the design of nanoparticles for optoelectronic applications. In this section cluster densities and energies will be discussed and radial distances will be used to consider the partitioning behaviour of clusters containing different ratios of molecule sizes.

Cluster diameters and densities show that cPAHs form tightly packed clusters. Homogeneous B clusters possess densities of 1.52–1.59 g/cm^3, which is comparable to mature graphitised soot particles (1.50–2.08 g/cm^3 [174]) although still significantly lower than graphite (2.09–2.23 g/cm^3). Cluster densities for A are between 1.45 and 1.50 g/cm^3, higher

than that of the A crystal structure (1.36 g/cm^3 [255]) and comparably sized fPAH clusters (1.39–1.46 g/cm^3 [64]) but still within the range ascribed to young soot particles (1.12–1.50 g/cm^3 [323, 51]). All cluster densities decrease with increasing cluster size, although this effect is muted for cPAH clusters compared to their fPAH counterparts. The diameter and density values of heterogeneous cPAH clusters are consistent with simple mixing averages of the analogous homogeneous cPAH clusters, suggesting that heterogeneity does not cause a dramatic change in the overall cluster shape and packing.

There is particular interest in whether cPAHs show the same core-shell partitioning seen in fPAH systems, where the cluster core consists of the larger fPAHs and the shell contains the smaller fPAHs, which has been suggested to explain the core-shell nanostructure seen in young combustion carbons. Average radial distances between each molecule type and the cluster centre are calculated to provide an indication of the molecule type positioning within each cluster, tabulated in Table 8.2. For the homogeneous cases, the two molecule types show similar relative average radial distances at 77–85% of the total cluster radius, similar to the fPAH clusters (for example, C_{100} has a value of 77% [64]). However, in all clusters containing two molecule types A possess larger average radial distances (located within 80–90% of the total cluster radius) than B (70–80% of the cluster radius). This is indicative of a core-shell structure in which the larger molecules reside in the cluster core, as seen in fPAH clusters (for example, the average radial distances of C and D are 89% and 65% of the cluster radius, respectively, for $C_{16}D_{16}$).

Although the average radial distances show a core-shell partitioning of molecule sizes, the distribution of radial distances shows that this radial separation is less distinct for clusters containing cPAHs compared to similar fPAH clusters. Figure 8.8 shows that there is significant mixing of the molecule types within the cPAH clusters and molecules of both types are present near the cluster centre, regardless of cluster size or molecular ratio. This core-shell structure confirms that like fPAH clusters, cPAH clusters show the inverse molecular arrangement compared to that observed experimentally for mature combustion particulates [41], indicating that the core-shell nanostructure arises from carbonisation rather than physical partitioning of different sized PAHs.

The contribution of flexoelectrically polarised aromatics to soot formation was explored previously [216] and it follows that long-range dipole-dipole interactions between cPAHs may allow cPAH clusters to possess increased stability compared to fPAHs. This would be of particular interest to contexts in which stable PAH nanoparticles are present, such as combustion and interstellar medium. To address this hypothesis, the intermolecular energies as a function of mass for cPAH and fPAH clusters are presented in Figure 8.9. The energies

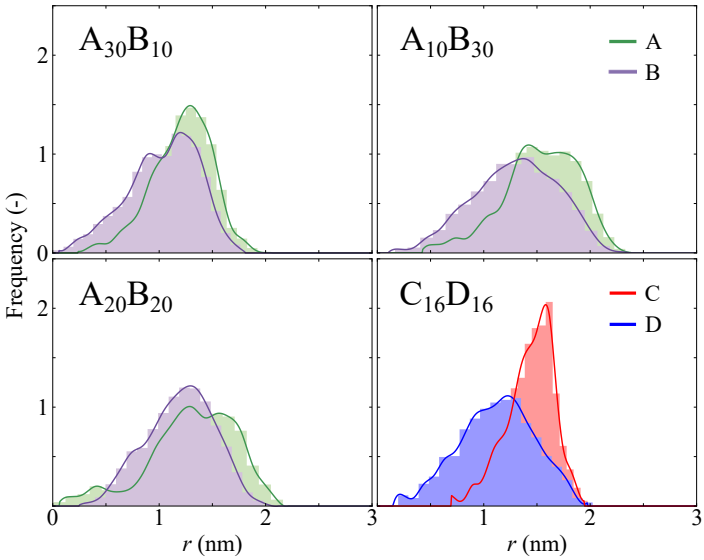


Fig. 8.8 Normalised atomic radial distance distributions for heterogeneous clusters containing curved and planar PAHs.

are shown per atom in the system in order to facilitate direct comparison between systems of different cluster and molecule sizes. It is clear that in all cases the energy decreases with cluster mass. Homogeneous clusters containing fPAHs (all coloured similarly here to allow for ease in reading), taken from Chen et al. [64, 61], show relatively consistent energy trends across molecule sizes from pyrene to circumcoronene, with heterogeneous clusters, taken from the work described in Chapter 4, generally at lower energies. Clusters containing cPAHs tend to have lower energies than fPAHs, as predicted based on their increased electrostatic attraction, although this is dependent on cPAH cluster composition. Clusters containing B show energies similar to those of heterogeneous fPAH clusters and A clusters possess lower mass-weighted energies. All cPAH clusters show significantly lower energies than the corresponding cPAH dimer interactions, which possess energies of -0.93 to -1.28 kJ/mol atom. For reference, the experimental binding energy of graphite is -3.4 kJ/mol atom with a large range from -2.4 to -4.8 kJ/mol atom [25].

The effect of heterogeneity in cPAH clusters shows a distinct effect of molecular ratio: an increased proportion of B in the cluster decreases the energy. When there are equal numbers of both molecules, the heterogeneous cluster energies reflect those of homogeneous A clusters across cluster sizes. However, changing the molecule proportions affects the cluster energies significantly in that $A_{30}B_{10}$ possesses a high energy while $A_{10}B_{30}$ has a significantly lower energy, producing the highest and lowest energy clusters of all the cPAH clusters evaluated. This suggests that in all cases the presence of B increases cluster stability if the composition is above 50%.

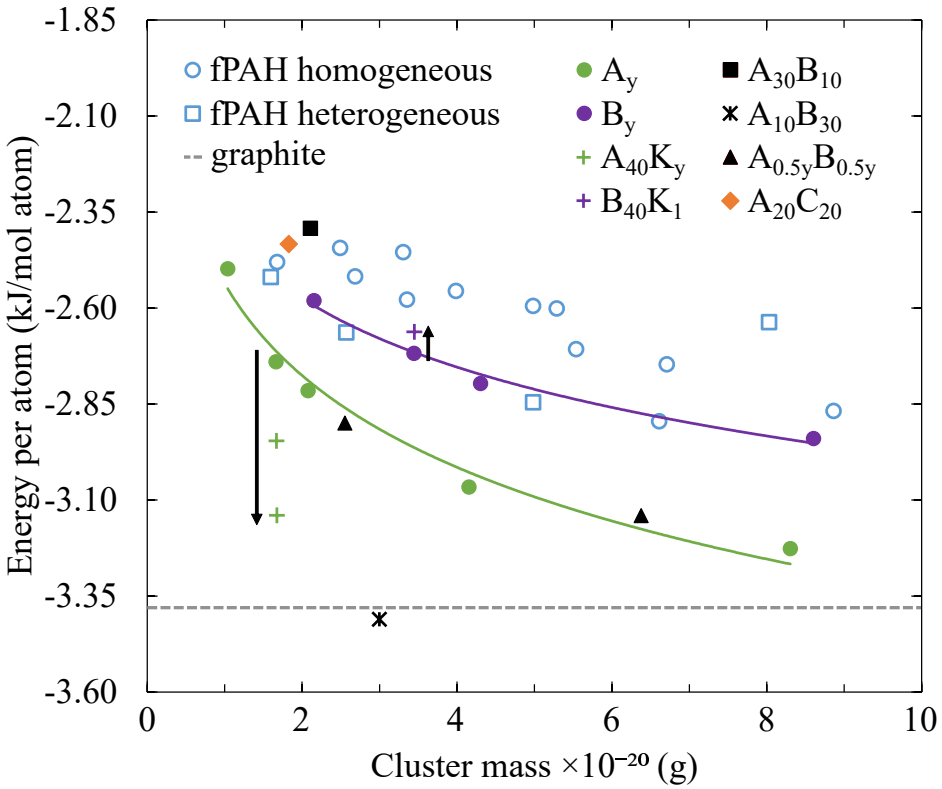


Fig. 8.9 Intermolecular energy per atom versus cluster mass for all clusters considered in this chapter. Lines are drawn for the cPAH homogeneous clusters to guide the eye and black arrows show the energy changes caused by the addition of a cation(s). The subscript y is used to denote an unspecified number of molecules in order to consider multiple clusters. The experimental binding energy of graphite [25] is shown as a horizontal dashed line.

These energy trends are in agreement with the structural metrics discussed previously, so that a higher degree of order within the heterogeneous cPAH clusters (increased stack formation as quantified by CNs, for example) corresponds to a greater interaction strength. This highlights that the stacking of A within the bowls of B decreases the cluster intermolecular energy until a proportion (>50% A) at which the presence of A disrupts B stacking by inserting within columns instead of only residing at column ends, thereby destabilising the cluster. These stable cPAH particles may present a potential nanostructure of soot particles but the enhanced cPAH-cPAH interactions are not sufficient to explain their rapid formation in flames.

8.3.3 How do complex cPAH systems self-assemble?

Many practical applications involve systems that contain many molecule types. Of particular interest to combustion and materials scientists are cPAH systems that also contain fPAHs

or ions. Relevant representative clusters are investigated using all of the metrics previously discussed to show the influence of different molecule curvatures and ion interactions within cPAH systems.

Particles containing cPAHs and fPAHs

To evaluate the self-assembly of clusters containing curved and flat PAHs, the cluster $A_{20}C_{20}$ is considered. As seen in Figure 8.3, this cluster shows a distinct partitioning of the molecule types in a janus configuration, in contrast to the mixed arrangement seen in homogeneous clusters and the core-shell structure seen in heterogeneous clusters containing either fPAHs or cPAHs. The average intermolecular spacing of 0.43 nm (Table 8.2) comes almost entirely from neighbouring C and matches the spacing calculated for a homogeneous C cluster [63]. This suggests that the fPAH molecules orient in a stacked configuration, but the constituent cPAHs do not. This is highlighted in a detailed CN evaluation which shows constituent C have a CN of approximately 1.5 while A have no near neighbours within a stacked configuration cut-off radius (CN of 0.0). These structural values agree with the respective homogeneous C and A clusters, as do the molecular alignment angles shown in Figure 8.10. The distinct types of local ordering cause the mixed fPAH and cPAH cluster to produce a janus particle with two different halves, highlighted by similar molecule type radial distances (see Table 8.2). This cluster containing both curved and flat PAHs has an energy similar to that of homogeneous cPAH clusters, indicating that the mixing of molecule types does not enhance molecule interaction energy and instead it remains similar to that of the weaker component.

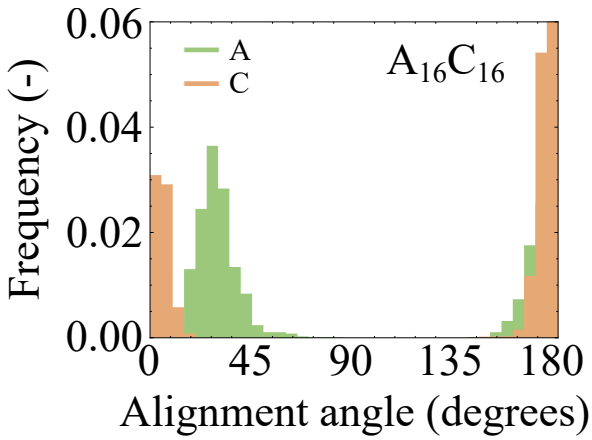


Fig. 8.10 Alignment angle distributions for the $A_{20}C_{20}$ cluster containing both cPAHs and fPAHs.

Dispersive and electrostatic interactions dominate the attractive forces between homogeneous interactions of fPAHs and cPAHs. However, this mixed molecule system is hindered by sterics and mismatched polarities and thus produces weaker interactions and a relatively low density. In this way the stronger interactions between like molecules contribute to a particle system in which the two molecule types are immiscible. Many material systems, such as coal and soot particles, are seen experimentally to possess both curved and flat aromatics. These results suggest that when molecule type separation is not observed in such systems, self-assembly through physical interactions does not control the molecular structure. For example, covalent bonding between fPAHs and cPAHs may explain why soot particles do not form janus particles [250, 251]. These results are also of interest in developing janus colloidal systems where interesting flow dynamics and self-assembly are predicted [369, 280].

Particles containing cPAHs and ion(s)

The influence of heterogeneity is further explored by considering clusters containing 40 molecules (A or B) and one or two potassium cations. These systems can be directly compared to homogeneous systems of the same size. Potassium ions are selected because they are known to interact with cPAHs in systems of interest [300] and, importantly, because the intermolecular interactions between K^+ and cPAHs have already been parametrised within the curPAHIP potential. As seen in Table 8.2, the inclusion of ion(s) within the cPAH clusters appears to increase the cluster diameter, and thus decrease the cluster density, slightly.

The cation(s) seems to have an effect on molecular arrangement that depends on the constituent cPAH size. The average CN of the B cluster decreases with the addition of K^+ (shown by a dotted arrow in Figure 8.4), suggesting a disruption of the highly stacked ordering in agreement with an increased intermolecular spacing. Conversely, the A clusters show increased ordering with a higher CN (dashed arrow in Figure 8.4) and decreased spacing. This effect increases with the addition of a second cation suggesting that the cation influence is additive, although the CNs are still lower than those of the B clusters. The addition of a potassium cation(s) also results in decreased average radial distances for both molecule types, so that average molecular distance values are around 75% of the total cluster radius. In agreement with these structural metrics, the presence of cation(s) in the cluster has opposite effects on cluster energies: B clusters with K^+ show an increased energy, while clusters with A containing K^+ showed a decreased energy in proportion with the number of ions. These differences are shown with black arrows in Figure 8.9, again showing that the presence of an ion(s) influences molecular interactions and cluster structure in different ways depending on the molecule size.

The majority of B within an ion-containing cluster remain in tilted stacks with the alignment angle distribution, seen in Figure 8.6 (bottom row), showing a dominant peak at 20°. Unlike the homogeneous B cluster, however, additional angles are present around 55° and 135°. This shows that the addition of a cation causes some disruption to the stacked structure otherwise seen in homogeneous or heterogeneous clusters containing B. This is visible in the cluster image in Figure 8.3, which shows that the $B_{40}K_1$ possesses more short stacks compared to the long columnar structure of B_{40}. The alignment angle distributions of A clusters containing K^+ are significantly different than the homogeneous and heterogeneous cases as well as the crystal structure, with a broadening and splitting of the 45° peak, shift and increase of a peak around 110°, and disappearance of high angles above 140°.

These quantitative differences suggest that the addition of a cation promotes stacking within the A cluster and disrupts the existing stacking within the B cluster. This can be observed in the cluster visualisations in Figure 8.3, which highlight the atoms immediately around the cation(s). All clusters show a solvation shell of four molecules around each cation with the cPAH electron-rich convex surfaces facing towards the cation(s). cPAHs form a tetrahedral-like formation, called a "flower" motif previously (Chapter 7), compared to the staggered triangular "propeller" motif adopted by fPAHs around a cation [21]. Steric effects allow this motif to serve as a seed for the A structure while disrupting the long stacks otherwise present in B clusters. This indicates that the self-assembly behaviour around cations depends on the cPAHs considered and so the addition of cations does not always increase molecular alignment and stability in large homogeneous clusters. These results also support the idea that a small cluster of 1–5 cPAH can be stabilised as an ionic nuclei that could contribute to other mechanisms for soot nanoparticle formation [216].

8.4 Conclusions

Curved carbons are ubiquitous and understanding such systems holds great promise in many systems and applications such as pollution reduction, novel carbon nanomaterials, and sensors. This chapter provides a first detailed exploration of the self-assembly of particles containing cPAHs, examining their nanostructure in homogeneous clusters containing one molecule type only, heterogeneous clusters containing different cPAH sizes and ratios, and complex cPAH clusters containing fPAHs or cations. The previously developed curPAHIP potential is extended to capture the interactions between large cPAHs and mixed cPAH systems and allow the simulations necessary for this analysis. Three questions in materials and combustion science are addressed, seeking to understand the ability of cPAHs to self-

assemble into an ordered phase, the characteristics of the internal nanostructure of cPAH nanoparticles, and the influence of fPAHs and ions within cPAH systems.

The presence of cPAHs generally disrupts mesophase formation, although the structures formed depend on the constituent cPAHs. Structural metrics, including intermolecular spacing, coordination number and alignment angles show that the nanostructure of homogeneous cPAH particles is dependent on molecule size but unaffected by particle size (from 2–5 nm). Clusters containing $C_{42}H_{14}$ (B) form parallel stacked columns similar to fPAHs but with a distinct tilt in the molecular alignment angles. In contrast, clusters containing $C_{20}H_{10}$ molecules (A) do not possess short-range order. Both homogeneous particle systems show similarities to the crystal structures of comparable cPAHs. Mixing of cPAH sizes influences the structure of a nanoparticle. The addition of A to a particle containing B serves to enhance the π-π stacking and interactions of A, resulting in lower energy particles. CN and energy values show that B clusters with some added A (<50%) form stable nanoparticles arranged in molecular stacks that have stronger intermolecular interactions than homogeneous and heterogeneous clusters containing cPAHs and/or fPAHs. Regardless of molecular ratio, heterogeneous cPAH particles show a core-shell type arrangement similar to the partitioning observed in fPAH particles but less distinct due to bowl complementarity. Particles containing cPAHs and fPAHs self-assemble into janus particles with minimal interaction between the molecule types, suggesting that the two molecule types are not miscible and may provide nanoparticles with interesting electronic properties. The addition of cations within cPAH particles causes the formation of a solvation shell that influences the internal particle structure. A show enhanced local order extending from the cation(s), while the stacked structure of B is disrupted due to the cation's presence.

This chapter provides a detailed evaluation of the molecular structure of cPAH-based nanoparticles. The systems evaluated are not exhaustive but serve as a guide for how composition, such as molecule curvature, size, ratio, type and number, influences nanoparticle self-assembly. This provides insight into the probable arrangement of curved aromatics within systems of interest to both ubiquitous natural materials and those developed for applications. Further work to thoroughly develop a fPAH-cPAH intermolecular potential and test the effect of molecular composition should be done to further explore the formation and potential applications of semi-polar janus particles containing cPAHs and fPAHs. In addition, future studies should examine the influence of ion charge, size and number on these structural properties. The presence of additional atoms within cPAHs should also be considered since the interactions between curved aromatic molecules containing heteroatoms, such as oxygen, show increased dimer interactions that provide further electrostatic stabilisation [49].

Chapter 9

Conclusions and future work

9.1 Summary

Investigations of the intermolecular clustering of PAH molecules in the context of soot particle formation and behaviour have been ongoing for a number of decades. However, previous studies have often been based on simple models that ignore any heterogeneity in the particle composition despite experimental evidence showing that this is a prominent feature of these systems. Very little attention has been paid to the structure and properties of young soot particles using high accuracy methods at atomic resolution. In addition, no detailed studies have explored how curved soot precursor molecules are able to dynamically interact and self-assemble into particles. The interactions and behaviours of molecules within carbonaceous particles are crucial to testing particle formation hypotheses and understanding the nature of these ubiquitous nanoparticles and so the uncertainty surrounding heterogeneous particle properties presents a barrier to the development of a detailed understanding of these systems. To address this, this thesis investigates the self-assembly and structure of carbonaceous nanoparticles in the context of combustion-produced soot particles through the development and application of molecular modelling tools. This thesis can be categorised as contributing to the extension of carbonaceous nanoparticle models to consider compositional heterogeneity and its effects on particle formation and structure, the exploration of the potential role of curved aromatic molecules in soot particle nucleation and carbon-based nanomaterials, and method development to enable the high-accuracy evaluation of these complex PAH systems. The main investigations and findings of each results chapter of the thesis are summarised below.

Replica exchange molecular dynamics simulations using the accurate isoPAHAP potential were executed to provide stable structures of low energy clusters containing planar PAHs of different sizes and proportions. An evaluation of these structures showed that, regardless of initial configuration and molecule types, heterogeneous clusters partition into a core-shell structure in which molecule sizes increase from the core to the shell. This is contrary to what was expected from experimental images of mature soot particles, suggesting that these microscope-observed arrangements are caused by chemical growth and interactions not captured in this physical model used. These computational studies were proposed to be representative of young soot particles that have not undergone significant carbonisation. These results guided further experimental work mapping the radial distribution of molecule sizes within nascent soot particles, which confirmed the predictions proposed here with similar particle morphologies.

The surface features and composition of heterogeneous PAH clusters are crucial properties that determine particle reactivity and growth/destruction. A detailed evaluation of heterogeneous soot particle surface properties was conducted using molecular modelling methods, focusing on the influence of particle size, molecular composition, and temperature. This showed for the first time the effect of molecular heterogeneity on the surface composition of PAH clusters and provided relationships between particle diameter, temperature, and composition with the particle morphology, particle roughness, and the availability of reactive sites on the particle surface. This highlighted structural properties of these systems and provided insight into parameters often used by combustion models to represent particle properties and behaviour.

A global optimisation method was developed to find minimum energy configurations of clusters containing a large number of planar PAH molecules. Previous studies using stochastic methods to evaluate these systems had been plagued by sampling issues stemming from the unphysical intersections of ringed molecules and were limited to small clusters. This new method, called the Sphere Encapsulated Monte Carlo (SEMC) method, incorporates sphere-based rearrangement and minimisation steps into the established basin-hopping global optimisation method. It was applied to a variety of PAH clusters and its usefulness was highlighted through the determination of low energy clusters with similar energetic and structural properties to those calculated using REMD at a significantly reduced cost. In addition, the SEMC method allowed the evaluation of large and more complex PAH clusters beyond the scope of existing methods.

Experimental and computational studies showing a significant presence of curved PAHs in combustion environments and their distinctive steric and electronic properties led to a hypoth-

esis that the physical interactions of curved aromatics play a role in soot particle nucleation and their presence influences particle nanostructure. To explore this using high-accuracy dynamic studies, an intermolecular potential called curPAHIP was developed. This potential extends the isoPAHAP potential for planar PAHs by including flexoelectric dipole moments and enhanced dispersion interactions between curved molecules with embedded pentagonal rings. The curPAHIP potential was applied within molecular dynamics simulations examining the clustering behaviour of curved and planar PAHs with and without cations. Curved PAHs and cations showed significant and rapid nucleation of small stable clusters compared to all other systems, suggesting that long-range electrostatic interactions may play a role in soot nucleation.

The nanostructure of clusters containing homogeneous and heterogeneous curved PAHs was evaluated using REMD simulations, focusing on the effects of molecule size and curvature, particle molecular composition, and interactions with planar PAHs or ions. This involved extending the curPAHIP potential to larger cPAHs and highlights its potential for further work examining flexoelectric effects in carbon materials. The simulations showed that molecule curvature provides a similar core-shell structure within nanoparticles as seen with planar PAHs. However, for small curved PAH sizes molecular alignment was disrupted, suggesting that the graphitisation of these carbon materials would be impeded. The presence of ions stabilised small clusters of curved PAHs and changed the cluster internal structure while no mixing took place between curved and planar PAHs within a cluster.

9.2 Suggested future work

There remain many important open questions regarding the mechanism of soot formation and the nanostructure of these small self-assembled particles. There are many avenues for potential future work, some building directly off the work in this thesis and others that go beyond the methods and systems addressed here.

9.2.1 Surface properties of clusters containing curved PAHs

Following the work presented in Chapters 5 and 8, a natural next step would be to apply solvent-excluded surface analysis to the clusters containing curved PAHs. This would provide further information about the detailed structure of nanoparticles relevant to combustion and may highlight features that make these systems useful to other applications.

9.2.2 Nucleation of large cPAHs and other ions

Another useful study would involve expanding the work presented in Chapter 7 to explore more complex systems relevant to soot formation. This could help refine an understanding regarding whether large curved PAH molecules are able to form stable clusters at higher temperatures than observed for corannulene. In addition, an analysis of the influence of ion size and charge (for example, comparing cations versus anions and/or studying chemi-ions) would further the understanding of nucleating interactions within the combustion process. The use of advanced free energy methods within molecular dynamics could provide further information on the nucleation cluster size and clustering free energy.

9.2.3 Heterogeneous cluster melting points

It may be possible to extract useful thermodynamic information from the multi-temperature trajectories of cPAH clusters such as those produced in Chapter 8. Using energy- and movement-based metrics, such as intermolecular energy changes and the Lindemann Index, melting and boiling points could be determined. This fundamental information would be useful for understanding nanoparticle behaviour in temperature-dependent processes, including soot particle nucleation and growth.

9.2.4 Extension to study curved carbon materials

Although focused on soot particles, many of the contributions in this thesis may have relevance to carbon materials in other contexts and further work should explore this. In particular, the charge polarisation found in carbon curved by pentagon integration has significant implications for the use of carbon materials in applications such as batteries (for example, ion batteries could be influenced by the charge polarisation of carbon electrodes) and adsorption (for example, both polar and non-polar adsorbents could interact strongly with flexoelectric materials). A rigorous evaluation of these interactions and their role in different contexts would involve the development of a more general form of the curPAHIP potential developed in Chapter 7 that incorporates polarisability and adaptable point charges for varied and complex carbon materials. This potential would provide a framework from which molecular dynamics and grand canonical Monte Carlo simulations could understand and optimise existing materials and propose new nanoengineered carbon materials.

9.2.5 Dynamic study of soot formation incorporating both physical and chemical interactions

The mechanism of soot formation is an area of ongoing investigation and it is increasingly clear that no single mechanism can fully account for the formation of soot particles. Recent work has suggested new mechanisms that combine both physical and chemical interactions between gas-phase molecules. Martin et al. [219] identified that PAHs possessing five-membered ring(s) on the molecule edge can be partially saturated localised π-radicals that possess a high reactivity and are found in reasonably high concentrations in the flames. These molecules are able to form covalent bonds and also physically condense in so-called "pancake bonds". This is supported by calculations by Gentile et al. [122]. Johansson et al. [175] have suggested that resonance-stabilised radical species may react with other hydrocarbon molecules to form and grow particles through covalent bonding. Frenklach and Mebel [119] has suggested that collisions between an aromatic molecule containing an edge five-membered ring and an aromatic radical are able to form a cross-link with the formation of rotationally-activated dimers. These proposed pathways show promise although further work is required to investigate these interactions in dynamic environments and identify the rates and identities of both precursor molecules and nucleated particles. Molecular dynamics provides a useful method to investigate these findings and provide insight into the mechanisms and rates of soot particle nucleation. This would require the parametrisation of a reactive force field that is able to capture both the physical and chemical bonding and energetics of relevant PAHs.

References

[1] Abbott, B. P., Abbott, R., Abbott, T., Abernathy, M., Acernese, F., Ackley, K., Adams, C., Adams, T., Addesso, P., Adhikari, R., et al. (2016). Observation of gravitational waves from a binary black hole merger. *Physical Review Letters*, 116(6):061102.

[2] Abdel-Shafy, H. I. and Mansour, M. S. (2016). A review on polycyclic aromatic hydrocarbons: source, environmental impact, effect on human health and remediation. *Egyptian Journal of Petroleum*, 25(1):107–123.

[3] Abid, A., Tolmachoff, E., Phares, D., Wang, H., Liu, Y., and Laskin, A. (2009). Size distribution and morphology of nascent soot in premixed ethylene flames with and without benzene doping. *Proceedings of the Combustion Institute*, 32(1):681–688.

[4] Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1:19–25.

[5] Abrahamson, J. P., Jain, A., van Duin, A. C., and Vander Wal, R. L. (2018). Carbon structure and the resulting graphitizability upon oxygen evolution. *Carbon*, 135:171–179.

[6] Adkins, E. M., Giaccai, J. A., and Miller, J. H. (2017). Computed electronic structure of polynuclear aromatic hydrocarbon agglomerates. *Proceedings of the Combustion Institute*, 36(1):957–964.

[7] Adkins, E. M. and Miller, J. H. (2015). Extinction measurements for optical band gap determination of soot in a series of nitrogen-diluted ethylene/air non-premixed flames. *Physical Chemistry Chemical Physics*, 17(4):2686–2695.

[8] Aguado, A., Vega, A., Lebon, A., and von Issendorff, B. (2018). Are zinc clusters really amorphous? A detailed protocol for locating global minimum structures of clusters. *Nanoscale*, 10(40):19162–19181.

[9] Akridis, P. and Rigopoulos, S. (2015). Modelling of soot formation in a laminar coflow non-premixed flame with a detailed CFD-population balance model. *Procedia Engineering*, 102:1274–1283.

[10] Alfè, M., Apicella, B., Barbella, R., Rouzaud, J. N., Tregrossi, A., and Ciajolo, A. (2009). Structure-property relationship in nanostructures of young and mature soot in premixed flames. *Proceedings of the Combustion Institute*, 32 I:697–704.

[11] Allendorf, H. and Rosner, D. (1968). Comparative studies of the attack of pyrolytic and isotropic graphite by atomic and molecular oxygen at high temperatures. *AIAA J.*, 6(4):650–654.

[12] Andersen, H. C. (1980). Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, 72(4):2384–2393.

[13] Antiñolo, M., Willis, M. D., Zhou, S., and Abbatt, J. P. (2015). Connecting the oxidation of soot to its redox cycling abilities. *Nature Communications*, 6(1):1–7.

[14] Apicella, B., Ciajolo, A., Tregrossi, A., Abrahamson, J., Vander Wal, R. L., and Russo, C. (2018). Hrtem and eels investigations of flame-formed soot nanostructure. *Fuel*, 225:218–224.

[15] Apicella, B., Pre, P., Alfè, M., Ciajolo, A., Gargiulo, V., Russo, C., Tregrossi, A., Deldique, D., and Rouzaud, J. (2015a). Soot nanostructure evolution in premixed flames by High Resolution Electron Transmission Microscopy (HRTEM). *Proceedings of the Combustion Institute*, 35(2):1895–1902.

[16] Apicella, B., Pré, P., Alfè, M., Ciajolo, A., Gargiulo, V., Russo, C., Tregrossi, A., Deldique, D., and Rouzaud, J. N. (2015b). Soot nanostructure evolution in premixed flames by High Resolution Electron Transmission Microscopy (HRTEM). *Proceedings of the Combustion Institute*, 35(2):1895–1902.

[17] Appel, J., Bockhorn, H., and Frenklach, M. (2000). Kinetic modeling of soot formation with detailed chemistry and physics: laminar premixed flames of C_2 hydrocarbons. *Combustion and Flame*, 121(1-2):122–136.

[18] Aubagnac-Karkar, D., El Bakali, A., and Desgroux, P. (2018). Soot particles inception and pah condensation modelling applied in a soot model utilizing a sectional method. *Combustion and Flame*, 189:190–206.

[19] Baletto, F. and Ferrando, R. (2005). Structural properties of nanoclusters: Energetic, thermodynamic, and kinetic effects. *Reviews of Modern Physics*, 77(1):371–423.

[20] Baquet, T. G., Grotheer, H.-H., and Aigner, M. (2007). Simultaneous detection of two types of soot precursor particles using photoionization mass spectrometry. *Rapid Communications in Mass Spectrometry*, 21(24):4060–4064.

[21] Bartolomei, M., Pirani, F., and Marques, J. (2019). Aggregation enhancement of coronene molecules by seeding with alkali-metal ions. *Physical Chemistry Chemical Physics*, 21(29):16005–16016.

[22] Bartolomei, M., Pirani, F., and Marques, J. M. C. (2017). Modeling Coronene Nanostructures: Analytical Potential, Stable Configurations and Ab Initio Energies. *The Journal of Physical Chemistry C*, 121:14330–14338.

[23] Bejaoui, S., Mercier, X., Desgroux, P., and Therssen, E. (2014). Laser induced fluorescence spectroscopy of aromatic species produced in atmospheric sooting flames using uv and visible excitation wavelengths. *Combustion and Flame*, 161(10):2479–2491.

[24] Benajes, J., Martín, J., García, A., Villalta, D., and Warey, A. (2015). In-cylinder soot radiation heat transfer in direct-injection diesel engines. *Energy Conversion and Management*, 106:414–427.

[25] Benedict, L. X., Chopra, N. G., Cohen, M. L., Zettl, A., Louie, S. G., and Crespi, V. H. (1998). Microscopic determination of the interlayer binding energy in graphite. *Chemical Physics Letters*, 286(5-6):490–496.

[26] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684–3690.

[27] Bernal, J. D. and Fowler, R. H. (1933). A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *The Journal of Chemical Physics*, 1(8):515–548.

[28] Bhowmick, H., Majumdar, S., and Biswas, S. (2012). Tribology of soot suspension in hexadecane as distinguished by the physical structure and chemistry of soot particles. *Journal of Physics D*, 45(17):175302.

[29] Biase, E. D. and Sarkisov, L. (2013). Systematic development of predictive molecular models of high surface area activated carbons for adsorption applications. *Carbon*, 64:262 – 280.

[30] Biswas, P., Wang, Y., and Attoui, M. (2018). Sub-2 nm particle measurement in high-temperature aerosol reactors: a review. *Current Opinion in Chemical Engineering*, 21:60–66.

[31] Blevins, L. G., Fletcher, R. A., Benner Jr, B. A., Steel, E. B., and Mulholland, G. W. (2002). The existence of young soot in the exhaust of inverse diffusion flames. *Proceedings of the Combustion Institute*, 29(2):2325–2333.

[32] Bockhorn, H., Andrea, D., Sarofim, A. F., and Wang, H., editors (2009). *Combustion Generated Fine Carbonaceous Particles*. KIT Scientific publishing.

[33] Boersma, C., Bregman, J. D., and Allamandola, L. J. (2013). Properties of polycyclic aromatic hydrocarbons in the northwest photon dominated region of NGC 7023. I. PAH size, charge, composition, and structure distribution. *Astrophysical Journal*, 769(2):1–13.

[34] Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, T., DeAngelo, B. J., Flanner, M. G., Ghan, S., Kärcher, B., Koch, D., et al. (2013). Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, 118(11):5380–5552.

[35] Bond, T. C. and Sun, H. (2005). Can reducing black carbon emissions counteract global warming? *Environmental Science & Technology*, 39:5921–5926.

[36] Bonne, U., Homann, K., and Wagner, H. G. (1965). Carbon formation in premixed flames. *Symposium (International) on Combustion*, 10(1):503–512.

[37] Bordat, P. and Brown, R. (2009). Structure and molecular dynamics of crystalline and liquid anthracene and naphthalene: Possible transient rotator phase of naphthalene. *The Journal of Chemical Physics*, 130(12):124501.

[38] Born, M. and Mayer, J. E. (1932). Zur gittertheorie der ionenkristalle. *Zeitschrift für Physik*, 75(1-2):1–18.

[39] Botero, M. L., Adkins, E. M., González-Calera, S., Miller, H., and Kraft, M. (2016a). PAH structure analysis of soot in a non-premixed flame using high-resolution transmission electron microscopy and optical band gap analysis. *Combustion and Flame*, 164:250–258.

[40] Botero, M. L., Chen, D., González-Calera, S., Jefferson, D., and Kraft, M. (2016b). HRTEM evaluation of soot particles produced by the non-premixed combustion of liquid fuels. *Carbon*, 96:459–473.

[41] Botero, M. L., Sheng, Y., Akroyd, J., Martin, J., Dreyer, J. A., Yang, W., and Kraft, M. (2019). Internal structure of soot particles in a diffusion flame. *Carbon*, 141:635–642.

[42] Bowal, K., Grančič, P., Martin, J. W., and Kraft, M. (2019a). Sphere Encapsulated Monte Carlo: Obtaining minimum energy configurations of large aromatic systems. *The Journal of Physical Chemistry A*, 123(33):7303–7313.

[43] Bowal, K., Martin, J. W., and Kraft, M. (2019b). Partitioning of polycyclic aromatic hydrocarbons in heterogeneous clusters. *Carbon*, 143:247–256.

[44] Bowal, K., Martin, J. W., Misquitta, A. J., and Kraft, M. (2019c). Ion-induced soot nucleation using a new potential for curved aromatics. *Combustion Science and Technology*, 191(5-6):747–765.

[45] Bowal, K., Pascazio, L., Wang, H., Chen, D., and Kraft, M. (2020). Surface properties of heterogeneous polycyclic aromatic hydrocarbon clusters. *Proceedings of the Combustion Institute*.

[46] Bronstein, H. E., Choi, N., and Scott, L. T. (2002). Practical synthesis of an open geodesic polyarene with a fullerene-type 6:6-double bond at the center: Diindeno[1,2,3,4-*defg*;1',2',3',4'-*mnop*]chrysene. *Journal of the American Chemical Society*, 124(30):8870–8875.

[47] Buckingham, R. A. (1938). The classical equation of state of gaseous helium, neon and argon. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 168(933):264–283.

[48] Byrd, R., Lu, P., Nocedal, J., and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal of Scientific Computing*, 16:1190–1208.

[49] Cabaleiro-Lago, E. M., Fernández, B., and Rodríguez-Otero, J. (2018). Dissecting the concave-convex π-π interaction in corannulene and sumanene dimers: SAPT(DFT) analysis and performance of DFT dispersion-corrected methods. *Journal of Computational Chemistry*, 39(2):93–104.

[50] Cain, J., Laskin, A., Kholghy, M. R., Thomson, M. J., and Wang, H. (2014). Molecular characterization of organic content of soot along the centerline of a coflow diffusion flame. *Physical Chemistry Chemical Physics*, 16(47):25862–25875.

[51] Camacho, J., Liu, C., Gu, C., Lin, H., Huang, Z., Tang, Q., You, X., Saggese, C., Li, Y., Jung, H., et al. (2015a). Mobility size and mass of nascent soot particles in a benchmark premixed ethylene flame. *Combustion and Flame*, 162(10):3810–3822.

[52] Camacho, J., Tao, Y., and Wang, H. (2015b). Kinetics of nascent soot oxidation by molecular oxygen in a flow reactor. *Proceedings of the Combustion Institute*, 35(2):1887–1894.

[53] Carbone, F., Canagaratna, M. R., Lambe, A. T., Jayne, J. T., Worsnop, D. R., and Gomez, A. (2018). Exploratory analysis of a sooting premixed flame via on-line high resolution (APi–TOF) mass spectrometry. *Proceedings of the Combustion Institute*.

[54] Carbone, F., Moslih, S., and Gomez, A. (2017). Probing gas-to-particle transition in a moderately sooting atmospheric pressure ethylene/air laminar premixed flame. Part II: Molecular clusters and nascent soot particle size distributions. *Combustion and Flame*, 181:329–341.

[55] Carrazana-García, J. A., Rodríguez-Otero, J., and Cabaleiro-Lago, E. M. (2011). DFT study of the interaction between alkaline cations and molecular bowls derived from fullerene. *Journal of Physical Chemistry B*, 115(12):2774–2782.

[56] Cauchy, A. (1847). Méthode générale pour la résolution des systemes d'équations simultanées. *Comp. Rend. Sci. Paris*, 25(1847):536–538.

[57] Celnik, M., Patterson, R., Kraft, M., and Wagner, W. (2007). Coupling a stochastic soot population balance to gas-phase chemistry using operator splitting. *Combustion and Flame*, 148(3):158–176.

[58] Chakrabarti, D., Fejer, S. N., and Wales, D. J. (2009). Rational design of helical architectures. *Proceedings of the National Academy of Sciences*, 106(48):20164–20167.

[59] Chen, D. (2014). *Numerical Investigation of Polycyclic Aromatic Hydrocarbon Clusters*. PhD thesis, University of Cambridge.

[60] Chen, D., Akroyd, J., Mosbach, S., and Kraft, M. (2015a). Surface reactivity of polycyclic aromatic hydrocarbon clusters. *Proceedings of the Combustion Institute*, 35(2):1811–1818.

[61] Chen, D., Akroyd, J., Mosbach, S., Opalka, D., and Kraft, M. (2015b). Solid-liquid transitions in homogenous ovalene, hexabenzocoronene and circumcoronene clusters: A molecular dynamics study. *Combustion and Flame*, 162(2):486–495.

[62] Chen, D. and Luo, K. H. (2020). Reactive sites on the surface of polycyclic aromatic hydrocarbon clusters: A numerical study. *Combustion and Flame*, 211:362–373.

[63] Chen, D., Totton, T. S., Akroyd, J., Mosbach, S., and Kraft, M. (2014a). Phase change of polycyclic aromatic hydrocarbon clusters by mass addition. *Carbon*, 77:25–35.

[64] Chen, D., Totton, T. S., Akroyd, J., Mosbach, S., and Kraft, M. (2014b). Size-dependent melting of polycyclic aromatic hydrocarbon nano-clusters: A molecular dynamics study. *Carbon*, 67:79–91.

[65] Chen, D. and Wang, H. (2017). Cation-π interactions between flame chemi-ions and aromatic compounds. *Energy & Fuels*, 31(3):2345–2352.

[66] Chen, M.-K., Hsin, H.-J., Wu, T.-C., Kang, B.-Y., Lee, Y.-W., Kuo, M.-Y., and Wu, Y.-T. (2014c). Highly curved bowl-shaped fragments of fullerenes: Synthesis, structural analysis, and physical properties. *Chemistry–A European Journal*, 20(2):598–608.

[67] Chernov, V., Thomson, M. J., Dworkin, S. B., Slavinskaya, N. A., and Riedel, U. (2014). Soot formation with c1 and c2 fuels using an improved chemical mechanism for pah growth. *Combustion and Flame*, 161(2):592–601.

[68] Chickos, J. S., Webb, P., Nichols, G., Kiyobayashi, T., Cheng, P.-C., and Scott, L. (2002). The enthalpy of vaporization and sublimation of corannulene, coronene, and perylene at T= 298.15 K. *The Journal of Chemical Thermodynamics*, 34(8):1195–1206.

[69] Chill, S. T., Welborn, M., Terrell, R., Zhang, L., Berthet, J.-C., Pedersen, A., Jonsson, H., and Henkelman, G. (2014). EON: Software for long time simulations of atomic scale systems. *Modelling and Simulation in Materials Science and Engineering*, 22(5):055002.

[70] Chodera, J. D., Swope, W. C., Pitera, J. W., Seok, C., and Dill, K. A. (2007). Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations. *Journal of Chemical Theory and Computation*, 3(1):26–41.

[71] Chung, S.-H. and Violi, A. (2011). Peri-condensed aromatics with aliphatic chains as key intermediates for the nucleation of aromatic hydrocarbons. *Proceedings of the Combustion Institute*, 33(1):693–700.

[72] Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., et al. (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the global burden of diseases study 2015. *The Lancet*, 389(10082):1907–1918.

[73] Contreras, J., Rodriguez, E., and Taha-Tijerina, J. (2017). Nanotechnology applications for electrical transformers—a review. *Electric Power Systems Research*, 143:573–584.

[74] Cozzini, S. and Ronchetti, M. (1996). Local icosahedral structures in binary-alloy clusters from molecular-dynamics simulation. *Physical Review B*, 53(18):12040–12049.

[75] D'Alessio, A., D'Anna, A., D'orsi, A., Minutolo, P., Barbella, R., and Ciajolo, A. (1992). Precursor formation and soot inception in premixed ethylene flames. *Symposium (International) on Combustion*, 24(1):973–980.

[76] D'Anna, A. (2009). Combustion-formed nanoparticles. *Proceedings of the Combustion Institute*, 32(1):593–613.

[77] D'Anna, A., Violi, A., D'Alessio, A., and Sarofim, A. (2001). A reaction pathway for nanoparticle formation in rich premixed flames. *Combustion and Flame*, 127(1-2):1995–2003.

[78] Dat, N.-D. and Chang, M. B. (2017). Review on characteristics of PAHs in atmosphere, anthropogenic sources and control technologies. *Science of the Total Environment*, 609:682–693.

[79] Davis, J., Tiwari, K., and Novosselov, I. (2019). Soot morphology and nanostructure in complex flame flow patterns via secondary particle surface growth. *Fuel*, 245:447–457.

[80] Demir, B. and Göktuğ Ahunbay, M. (2016). Adsorption of perfluorohexane in BAM-P109 type activated carbon via molecular simulation. *Adsorption Science & Technology*, 34(1):79–92.

[81] Di Stasio, S. (2001). Electron microscopy evidence of aggregation under three different size scales for soot nanoparticles in flame. *Carbon*, 39(1):109–118.

[82] Di Stasio, S., Legarrec, J. L., and Mitchell, J. B. (2011). Synchrotron radiation studies of additives in combustion, II: Soot agglomerate microstructure change by alkali and alkaline-earth metal addition to a partially premixed flame. *Energy & Fuels*, 25(3):916–925.

[83] Di Stasio, S., Mitchell, J., LeGarrec, J.-L., Biennier, L., and Wulff, M. (2006). Synchrotron saxs <in situ> identification of three different size modes for soot nanoparticles in a diffusion flame. *Carbon*, 44(7):1267–1279.

[84] Dobbins, R. and Subramaniasivam, H. (1994). *Soot precursor particles in flames*, pages 280–301. Springer Science & Business Media.

[85] Dobbins, R. A. (2002). Soot inception temperature and the carbonization rate of precursor particles. *Combustion and Flame*, 130(3):204–214.

[86] Dobbins, R. A. (2007). Hydrocarbon nanoparticles formed in flames and diesel engines. *Aerosol Science and Technology*, 41(5):485–496.

[87] Dobbins, R. A., Fletcher, R. A., and Chang, H.-C. (1998). The evolution of soot precursor particles in a diffusion flame. *Combustion and Flame*, 115(3):285–298.

[88] Dobrowolski, M. A., Ciesielski, A., and Cyrański, M. K. (2011). On the aromatic stabilization of corannulene and coronene. *Physical Chemistry Chemical Physics*, 13(46):20557–20563.

[89] Doye, J. P. (2006). Lead clusters: Different potentials, different structures. *Computational Materials Science*, 35(3):227–231.

[90] Doye, J. P. and Meyer, L. (2005a). Mapping the magic numbers in binary Lennard-Jones clusters. *Physical Review Letters*, 95(6):063401.

[91] Doye, J. P., Miller, M. A., and Wales, D. J. (1999). Evolution of the potential energy surface with size for Lennard-Jones clusters. *The Journal of Chemical Physics*, 111(18):8417–8428.

[92] Doye, J. P. and Wales, D. J. (2001). Polytetrahedral clusters. *Physical Review Letters*, 86(25):5719–5722.

[93] Doye, J. P., Wales, D. J., Branz, W., and Calvo, F. (2001). Modeling the structure of clusters of C_{60} molecules. *Physical Review B*, 64(23):235409.

[94] Doye, J. P. K. and Meyer, L. (2005b). Mapping the magic numbers in binary Lennard-Jones clusters. *Physical Review Letters*, 95(6):063401.

[95] Doye, J. P. K. and Meyer, L. (2006). The structure of binary Lennard-Jones clusters: The effects of atomic size ratio. *ArXiv Condensed Matter e-prints*, cond-mat/0604250:1–13.

[96] Dreyer, J. A., Poli, M., Eaves, N. A., Botero, M. L., Akroyd, J., Mosbach, S., and Kraft, M. (2019). Evolution of the soot particle size distribution along the centreline of an n-heptane/toluene co-flow diffusion flame. *Combustion and Flame*, 209:256–266.

[97] Dubceac, C., Sevryugina, Y., Kuvychko, I. V., Boltalina, O. V., Strauss, S. H., and Petrukhina, M. A. (2018). Self-assembly of aligned hybrid one-dimensional stacks from two complementary π-bowls. *Crystal Growth & Design*, 18(1):307–311.

[98] Eaves, N. A., Dworkin, S. B., and Thomson, M. J. (2015). The importance of reversibility in modeling soot nucleation and condensation processes. *Proceedings of the Combustion Institute*, 35(2):1787–1794.

[99] Eaves, N. A., Veshkini, A., Riese, C., Zhang, Q., Dworkin, S. B., and Thomson, M. J. (2012). A numerical study of high pressure, laminar, sooting, ethane–air coflow diffusion flames. *Combustion and Flame*, 159(10):3179–3190.

[100] Ebert, L. B., Scanlon, J. C., and Clausen, C. A. (1988). Combustion tube soot from a diesel fuel/air mixture: Issues in structure and reactivity. *Energy & Fuels*, 2(4):438–445.

[101] Elvati, P., Turrentine, K., and Violi, A. (2019). The role of molecular properties on the dimerization of aromatic compounds. *Proceedings of the Combustion Institute*, 37(1):1099–1105.

[102] Elvati, P. and Violi, A. (2013). Thermodynamics of poly-aromatic hydrocarbon clustering and the effects of substituted aliphatic chains. *Proceedings of the Combustion Institute*, 34(1):1837–1843.

[103] Experts, I. (2019). Carbon black - a global market overview. Technical report, Industry Experts.

[104] Fejer, S. N., Chakrabarti, D., and Wales, D. J. (2011). Self-Assembly of anisotropic particles. *Soft Matter*, 7(7):3553–3564.

[105] Fejer, S. N. and Wales, D. J. (2007). Helix self-assembly from anisotropic molecules. *Physical Review Letters*, 99(8):086106.

[106] Ferenczy, G., Reynolds, C., Winn, P., and Stone, A. (1998). MULFIT: a program for calculating electrostatic potential-fitted charges. *May be obtained by contacting AJ Stone, email address: ajs1@ cam. ac. uk*.

[107] Fernandez-Alos, V., Watson, J. K., vander Wal, R., and Mathews, J. P. (2011). Soot and char molecular representations generated directly from HRTEM lattice fringe images using Fringe3D. *Combustion and Flame*, 158(9):1807–1813.

[108] Ferrenberg, A. M. and Swendsen, R. H. (1989). Optimized Monte Carlo data analysis. *Computers in Physics*, 3(5):101–104.

[109] Filatov, A. S. and Petrukhina, M. A. (2010). Probing the binding sites and coordination limits of buckybowls in a solvent-free environment: Experimental and theoretical assessment. *Coordination Chemistry Reviews*, 254(17-18):2234–2246.

[110] Filatov, A. S., Scott, L. T., and Petrukhina, M. A. (2010). π-π interactions and solid state packing trends of polycyclic aromatic bowls in the indenocorannulene family: Predicting potentially useful bulk properties. *Crystal Growth and Design*, 10(10):4607–4621.

[111] Forkey, D. M., Attar, S., Noll, B. C., Koerner, R., Olmstead, M. M., and Balch, A. L. (1997). Crystallographic characterization of the molecular structure and solid state packing of the fullerene-shaped hydrocarbon c36h12. *Journal of the American Chemical Society*, 119(24):5766–5767.

[112] Frenkel, D. and Smit, B. (2001). *Understanding molecular simulation: From algorithms to applications*, volume 1. Academic Press.

[113] Frenklach, M. (1996). On surface growth mechanism of soot particles. *Symposium (International) on Combustion*, 26(2):2285–2293.

[114] Frenklach, M. (2002a). Method of moments with interpolative closure. *Chemical Engineering Science*, 57(12):2229–2239.

[115] Frenklach, M. (2002b). Reaction mechanism of soot formation in flames. *Physical Chemistry Chemical Physics*, 4(11):2028–2037.

[116] Frenklach, M. (2019). New form for reduced modeling of soot oxidation: Accounting for multi-site kinetics and surface reactivity. *Combustion and Flame*, 201:148–159.

[117] Frenklach, M. and Harris, S. J. (1987). Aerosol dynamics modeling using the method of moments. *Journal of Colloid and Interface Science*, 118(1):252–261.

[118] Frenklach, M., Liu, Z., Singh, R. I., Galimova, G. R., Azyazov, V. N., and Mebel, A. M. (2018). Detailed, sterically-resolved modeling of soot oxidation: Role of O atoms, interplay with particle nanostructure, and emergence of inner particle burning. *Combustion and Flame*, 188:284–306.

[119] Frenklach, M. and Mebel, A. M. (2020). On the mechanism of soot nucleation. *Physical Chemistry Chemical Physics*, 22(9):5314–5331.

[120] Frenklach, M. and Wang, H. (1991). Detailed modeling of soot particle nucleation and growth. *Symposium (International) on Combustion*, 23(1):1559–1566.

[121] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota,

K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, Jr., J. A., Peralta, J. E., Ogliaro, F., Bearpark, M. J., Heyd, J. J., Brothers, E. N., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A. P., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B., and Fox, D. J. (2016). Gaussian 16 Revision A.03. Gaussian Inc. Wallingford CT.

[122] Gentile, F. S., Picca, F., De Falco, G., Commodo, M., Minutolo, P., Causà, M., and D'Anna, A. (2020). Soot inception: A DFT study of σ and π dimerization of resonantly stabilized aromatic radicals. *Fuel*, 279:118491.

[123] Girifalco, L. A., Hodak, M., and Lee, R. S. (2000). Carbon nanotubes, buckyballs, ropes, and a universal graphitic potential. *Physical Review B*, 62:13104–13110.

[124] Glassman, I. (1989). Soot formation in combustion processes. *Symposium (International) on Combustion*, 22(1):295–311.

[125] Glassman, I., Yetter, R. A., and Glumac, N. G. (2014). *Combustion*. Elsevier.

[126] Gobre, V. V. and Tkatchenko, A. (2013). Scaling laws for van der Waals interactions in nanostructured materials. *Nature Communications*, 4:2341.

[127] Goedecker, S. (2004). Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *The Journal of Chemical Physics*, 120(21):9911–9917.

[128] Gonzalez, C. and Lim, E. C. (2000). A quantum chemistry study of the van der Waals dimers of benzene, naphthalene, and anthracene: Crossed (D_{2d}) and parallel-displaced (c_{2h}) dimers of very similar energies in the linear polyacenes. *The Journal of Physical Chemistry A*, 104(13):2953–2957.

[129] Gonzalez, C. and Lim, E. C. (2003). Evaluation of the Hartree-Fock Dispersion (HFD) model as a practical tool for probing intermolecular potentials of small aromatic clusters: Comparison of the HFD and MP2 intermolecular potentials. *Journal of Physical Chemistry A*, 107(47):10105–10110.

[130] Grabowsky, S., Weber, M., Chen, Y.-S., Lentz, D., Schmidt, B. M., Hesse, M., and Luger, P. (2010). Electron density of corannulene from synchrotron data at 12 K, comparison with fullerenes. *Zeitschrift für Naturforschung B*, 65(4):452–460.

[131] Grančič, P., Bylsma, R., Meekes, H., and Cuppen, H. M. (2015). Evaluation of all-atom force fields for anthracene crystal growth. *Crystal Growth & Design*, 15(4):1625–1633.

[132] Grančič, P., Martin, J. W., Chen, D., Mosbach, S., and Kraft, M. (2016). Can nascent soot particles burn from the inside? *Carbon*, 109:608–615.

[133] Grimme, S. (2006). Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry*, 27(15):1787–1799.

[134] Grotheer, H.-H., Wolf, K., and Hoffmann, K. (2011). Photoionization mass spectrometry for the investigation ofácombustion generated nascent nanoparticles and their relation toálaser induced incandescence. *Applied Physics B*, 104(2):367–383.

[135] Guarnieri, M. and Balmes, J. R. (2014). Outdoor air pollution and asthma. *The Lancet*, 383(9928):1581–1592.

[136] Han, S., Li, X., Nie, F., Zheng, M., Liu, X., and Guo, L. (2017). Revealing the initial chemistry of soot nanoparticle formation by ReaxFF molecular dynamics simulations. *Energy & Fuels*, 31(8):8434–8444.

[137] Hansen, J. and Nazarenko, L. (2004). Soot climate forcing via snow and ice albedos. *Proceedings of the National Academy of Sciences*, 101(2):423–428.

[138] Hansmann, U. H. and Wille, L. T. (2002). Global optimization by energy landscape paving. *Physical Review Letters*, 88(6):068105.

[139] Hanson, J. C. and Nordman, C. (1976). The crystal and molecular structure of corannulene, c20h10. *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry*, 32(4):1147–1153.

[140] Happold, J., Grotheer, H.-H., and Aigner, M. (2009). *Soot precursors consisting of stacked pericondensed PAHs*, pages 277–288. In [32].

[141] Harris, P. J. (2005). New perspectives on the structure of graphitic carbons. *Critical Reviews in Solid State and Materials Sciences*, 30(4):235–253.

[142] Harris, P. J. F. (2004). Fullerene-related structure of commercial glassy carbons. *Philosophical Magazine*, 84(29):3159–3167.

[143] Harris, P. J. F., Liu, Z., and Suenaga, K. (2008). Imaging the atomic structure of activated carbon. *Journal of Physics Condensed Matter*, 20(36).

[144] Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.

[145] Haynes, B., Jander, H., and Wagner, H. G. (1979). The effect of metal additives on the formation of soot in premixed flames. *Symposium (international) on Combustion*, 17(1):1365–1374.

[146] Haynes, B. S. and Wagner, H. G. (1981). Soot formation. *Progress in energy and combustion science*, 7(4):229–273.

[147] Heal, M. R., Kumar, P., and Harrison, R. M. (2012). Particles, air quality, policy and health. *Chemical Society Reviews*, 41(19):6606–6630.

[148] Herdman, J. D., Connelly, B. C., Smooke, M. D., Long, M. B., and Miller, J. H. (2011). A comparison of raman signatures and laser-induced incandescence with direct numerical simulation of soot growth in non-premixed ethylene/air flames. *Carbon*, 49(15):5298–5311.

[149] Herdman, J. D. and Miller, J. H. (2008). Intermolecular potential calculations for polynuclear aromatic hydrocarbon clusters. *The Journal of Physical Chemistry A*, 112(28):6249–6256.

[150] Hernández-Rojas, J., Bretón, J., Gomez Llorente, J., and Wales, D. (2004). Lowest-energy structures of $(C_{60})_n$ *X* X = Li^+, Na^+, K^+, Cl^-) and $(C_{60})_n$ *Y*Cl(Y = Li, Na, K) clusters for $n \leq 13$. *The Journal of Chemical Physics*, 121(24):12315–12322.

[151] Hernández-Rojas, J. and Calvo, F. (2019). Coarse-grained modeling of the nucleation of polycyclic aromatic hydrocarbons into soot precursors. *Physical Chemistry Chemical Physics*.

[152] Hernández-Rojas, J., Calvo, F., Niblett, S., and Wales, D. (2017). Dynamics and thermodynamics of the coronene octamer described by coarse-grained potentials. *Physical Chemistry Chemical Physics*, 19(3):1884–1895.

[153] Hernández-Rojas, J., Calvo, F., and Wales, D. J. (2016). Coarse-graining the structure of polycyclic aromatic hydrocarbons clusters. *Physical Chemistry Chemical Physics*, 18(20):13736–13740.

[154] Hesselmann, A., Jansen, G., and Schütz, M. (2005). Density-functional theory-symmetry-adapted intermolecular perturbation theory with density fitting: A new efficient method to study intermolecular interaction energies. *Journal of Chemical Physics*, 122:014103.

[155] Hestenes, M. R., Stiefel, E., et al. (1952). Methods of conjugate gradients for solving linear systems. *Journal of research of the National Bureau of Standards*, 49(6):409–436.

[156] Highwood, E. J. and Kinnersley, R. P. (2006). When smoke gets in our eyes: The multiple impacts of atmospheric black carbon on climate, air quality and health. *Environment International*, 32(4):560–566.

[157] Hobza, P., Selzle, H. L., and Schlag, E. W. (1996). Potential energy surface for the benzene dimer. results of ab initio CCSD (T) calculations show two nearly isoenergetic structures: T-shaped and parallel-displaced. *The Journal of Physical Chemistry*, 100(48):18790–18794.

[158] Homann, K. H. and Wagner, H. G. (1967). Some new aspects of the mechanism of carbon formation in premixed flames. *Symposium (International) on Combustion*, 11(1):371–379.

[159] Hoover, W. G. (1985). Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31:1695–1697.

[160] Hu, L., Wang, S., Zhang, B., and Zeng, Y. (2006). Structural changes in soot particles induced by diode laser irradiation. *Carbon*, 44(9):1725–1729.

[161] Hukushima, K. and Nemoto, K. (1996). Exchange Monte Carlo method and application to spin glass simulations. *Journal of the Physical Society of Japan*, 65(6):1604–1608.

[162] Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1):33–38.

[163] Hurt, R. H., Crawford, G. P., and Shim, H.-S. (2000). Equilibrium nanostructure of primary soot particles. *Proceedings of the Combustion Institute*, 28(2):2539–2546.

[164] Hussain, K., Hoque, R. R., Balachandran, S., Medhi, S., Idris, M. G., Rahman, M., and Hussain, F. L. (2018). Monitoring and risk analysis of pahs in the environment. *Handbook of Environmental Materials Management*, pages 1–35.

[165] Iavarone, S., Pascazio, L., Sirignano, M., De Candia, A., Fierro, A., de Arcangelis, L., and D'Anna, A. (2016). Molecular dynamics simulations of incipient carbonaceous nanoparticle formation at flame conditions. *Combustion Theory and Modelling*, 7830:1–13.

[166] Iijima, S. (1980). Direct observation of the tetrahedral bonding in graphitized carbon black by high resolution electron microscopy. *Journal of Crystal Growth*, 50(3):675–683.

[167] Imamura, K., Takimiya, K., Otsubo, T., and Aso, Y. (1999). Triphenyleno[1,12-*bcd*: 4,5-*b*'c'd':8,9-*b*"c"d"]trithiophene: the first bowl-shaped heteroaromatic. *Chemical Communications*, 18:1859–1860.

[168] Ishiguro, T., Takatori, Y., and Akihama, K. (1997). Microstructure of diesel soot particles probed by electron microscopy: First observation of inner core and outer shell. *Combustion and Flame*, 108(1):231–234.

[169] Iwamatsu, M. and Okabe, Y. (2004). Basin hopping with occasional jumping. *Chemical Physics Letters*, 399(4-6):396–400.

[170] Janowski, T., Ford, A. R., and Pulay, P. (2010). Accurate correlated calculation of the intermolecular potential surface in the coronene dimer. *Molecular Physics*, 108(3-4):249–257.

[171] Janowski, T., Pulay, P., Karunarathna, A. S., Sygula, A., and Saebø, S. (2011). Convex–concave stacking of curved conjugated networks: Benchmark calculations on the corannulene dimer. *Chemical Physics Letters*, 512(4-6):155–160.

[172] Jensen, K. P. and Jorgensen, W. L. (2006). Halide, ammonium, and alkali metal ion parameters for modeling aqueous solutions. *Journal of Chemical Theory and Computation*, 2(6):1499–1509.

[173] Jia, J., Bi, C., Zhang, J. J., and Chen, Z. (2019). Atmospheric deposition and vegetable uptake of polycyclic aromatic hydrocarbons (pahs) based on experimental and computational simulations. *Atmospheric Environment*, 204:135–141.

[174] Johansson, K., El Gabaly, F., Schrader, P., Campbell, M., and Michelsen, H. (2017a). Evolution of maturity levels of the particle surface and bulk during soot growth and oxidation in a flame. *Aerosol Science and Technology*, 51(12):1333–1344.

[175] Johansson, K., Head-Gordon, M., Schrader, P., Wilson, K., and Michelsen, H. (2018). Resonance-stabilized hydrocarbon-radical chain reactions may explain soot inception and growth. *Science*, 361(6406):997–1000.

[176] Johansson, K. O., Dillstrom, T., Elvati, P., Campbell, M. F., Schrader, P. E., Popolan-Vaida, D. M., Richards-Henderson, N. K., Wilson, K. R., Violi, A., and Michelsen, H. A. (2017b). Radical-radical reactions, pyrene nucleation, and incipient soot formation in combustion. *Proceedings of the Combustion Institute*, 36(1):799–806.

[177] Jones, J. E. (1924). On the determination of molecular fields.—ii. from the equation of state of a gas. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 106(738):463–477.

[178] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935.

[179] Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L. (2001). Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *The Journal of Physical Chemistry B*, 105(28):6474–6487.

[180] Kanao, E., Kubo, T., Naito, T., Matsumoto, T., Sano, T., Yan, M., and Otsuka, K. (2018). Differentiating π interactions by constructing concave/convex surfaces using a bucky bowl molecule, corannulene in liquid chromatography. *Analytical chemistry*, 91(3):2439–2446.

[181] Kazakov, A., Wang, H., and Frenklach, M. (1995). Detailed modeling of soot formation in laminar premixed ethylene flames at a pressure of 10 bar. *Combustion and Flame*, 100(1-2):111–120.

[182] Kennedy, I. M. (1997). Models of soot formation and oxidation. *Progress in Energy and Combustion Science*, 23(2):95–132.

[183] Kennedy, M. R., Burns, L. A., and Sherrill, C. D. (2012). Buckyplates and buckybowls: Examining the effects of curvature on π–π interactions. *The Journal of Physical Chemistry A*, 116(48):11920–11926.

[184] Khanna, R., Sahajwalla, V., and Hurt, R. H. (2005). An atomistic technique for simulating non-covalent interactions in large ensembles of high-molecular-weight polyaromatics. *Carbon*, 43(1):67 – 77.

[185] Kholghy, M. R., Afarin, Y., Sediako, A. D., Barba, J., Lapuerta, M., Chu, C., Weingarten, J., Borshanpour, B., Chernov, V., and Thomson, M. J. (2017a). Comparison of multiple diagnostic techniques to study soot formation and morphology in a diffusion flame. *Combustion and Flame*, 176:567–583.

[186] Kholghy, M. R., Kelesidis, G. A., and Pratsinis, S. E. (2018). Reactive polycyclic aromatic hydrocarbon dimerization drives soot nucleation. *Physical Chemistry Chemical Physics*, 20(16):10925–10938.

[187] Kholghy, M. R., Veshkini, A., and Thomson, M. J. (2016). The core-shell internal nanostructure of soot - A criterion to model soot maturity. *Carbon*, 100:508–536.

[188] Kholghy, M. R., Weingarten, J., Sediako, A. D., Barba, J., Lapuerta, M., and Thomson, M. J. (2017b). Structural effects of biodiesel on soot formation in a laminar coflow diffusion flame. *Proceedings of the Combustion Institute*, 36(1):1321–1328.

[189] Kofke, D. A. (2002). On the acceptance probability of replica-exchange Monte Carlo trials. *The Journal of Chemical Physics*, 117(15):6911–6914.

[190] Kone, A. and Kofke, D. A. (2005). Selection of temperature intervals for parallel-tempering simulations. *The Journal of Chemical Physics*, 122(20):206101.

[191] Kuffner, J. J. (2004). Effective sampling and distance metrics for 3D rigid body path planning. In *IEEE International Conference on Robotics and Automation*, pages 3993–3998.

[192] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*, 13(8):1011–1021.

[193] Kusumaatmaja, H., Whittleston, C. S., and Wales, D. J. (2012). A local rigid body framework for global optimization of biomolecules. *Journal of Chemical Theory and Computation*, 8(12):5159–5165.

[194] Lafleur, A. L., Howard, J. B., Marr, J. A., and Yadav, T. (1993). Proposed fullerene precursor corannulene identified in flames both in the presence and absence of fullerene production. *The Journal of Physical Chemistry*, 97(51):13539–13543.

[195] Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. *Nature*, 409:860–921.

[196] Leach, A. R. (2001). *Molecular modelling: Principles and applications*. Pearson Education.

[197] Lee, E. C., Kim, D., Jurečka, P., Tarakeshwar, P., Hobza, P., and Kim, K. S. (2007). Understanding of assembly phenomena by aromatic-aromatic interactions: Benzene Dimer and the substituted systems. *Journal of Physical Chemistry A*, 111(18):3446–3457.

[198] Lee, Y., Lamarque, J., Flanner, M., Jiao, C., Shindell, D., Berntsen, T., Bisiaux, M., Cao, J., Collins, W., Curran, M., et al. (2013). Evaluation of preindustrial to present-day black carbon and its albedo forcing from Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP). *Atmospheric Chemistry and Physics*, 13(5):2607–2634.

[199] Leon, G., Eaves, N., Akroyd, J., Mosbach, S., and Kraft, M. (2019). A new methodology to calculate process rates in a kinetic monte carlo model of pah growth. *Combustion and Flame*, 209:133–143.

[200] Lewes, V. B. (1894). Ii. the action of heat upon ethylene. *Proceedings of the Royal Society of London*, 55(331-335):90–107.

[201] Li, J., Zhu, X., Deng, S., Liu, X., Deng, Q., Yu, K., Sun, H., Zhang, X., Guo, H., Chen, W., et al. (2015). Exposure to polycyclic aromatic hydrocarbons and accelerated dna methylation ageing: an observational study. *The Lancet*, 386:S20.

[202] Li, Z., Lu, C., Xia, Z., Zhou, Y., and Luo, Z. (2007). X-ray diffraction patterns of graphite and turbostratic carbon. *Carbon*, 45(8):1686–1695.

[203] Li, Z. and Scheraga, H. A. (1987). Monte Carlo-minimization approach to the multiple-minima problem in protein folding. *Proceedings of the National Academy of Sciences of the United States of America*, 84(19):6611–6615.

[204] Lippmann, M. (2014). Toxicological and epidemiological studies of cardiovascular effects of ambient air fine particulate matter (pm2.5) and its chemical components:coherence and public health implications. *Critical Reviews in Toxicology*, 44(4):299–347.

[205] Liu, A. and Rigopoulos, S. (2019). A conservative method for numerical solution of the population balance equation, and application to soot formation. *Combustion and Flame*, 205:506–521.

[206] Liu, C., Singh, A. V., Saggese, C., Tang, Q., Chen, D., Wan, K., Vinciguerra, M., Commodo, M., De Falco, G., Minutolo, P., et al. (2019a). Flame-formed carbon nanoparticles exhibit quantum dot behaviors. *Proceedings of the National Academy of Sciences*, 116(26):12692–12697.

[207] Liu, M., Artyukhov, V. I., Lee, H., Xu, F., and Yakobson, B. I. (2013). Carbyne from first principles: chain of c atoms, a nanorod or a nanorope. *ACS Nano*, 7(11):10075–10082.

[208] Liu, P., Li, Z., Bennett, A., Lin, H., Sarathy, S. M., and Roberts, W. L. (2019b). The site effect on pahs formation in haca-based mass growth process. *Combustion and Flame*, 199:54–68.

[209] Llanio-Trujillo, J., Marques, J., and Pereira, F. (2011). An evolutionary algorithm for the global optimization of molecular clusters: Application to water, benzene, and benzene cation. *The Journal of Physical Chemistry A*, 115(11):2130–2138.

[210] Lovas, F. J., McMahon, R. J., Grabow, J. U., Schnell, M., Mack, J., Scott, L. T., and Kuczkowski, R. L. (2005). Interstellar chemistry: A strategy for detecting polycyclic aromatic hydrocarbons in space. *Journal of the American Chemical Society*, 127(12):4345–4349.

[211] Lowe, J. S., Lai, J. Y., Elvati, P., and Violi, A. (2015). Towards a predictive model for polycyclic aromatic hydrocarbon dimerization propensity. *Proceedings of the Combustion Institute*, 35(2):1827–1832.

[212] Lu, T. and Law, C. K. (2009). Toward accommodating realistic fuel chemistry in large-scale computations. *Progress in Energy and Combustion Science*, 35(2):192–215.

[213] Mao, Q., Hou, D., Luo, K. H., and You, X. (2018). Dimerization of polycyclic aromatic hydrocarbon molecules and radicals under flame conditions. *The Journal of Physical Chemistry A*, 122(44):8701–8708.

[214] Mao, Q., van Duin, A. C., and Luo, K. (2017). Formation of incipient soot particles from polycyclic aromatic hydrocarbons: a ReaxFF molecular dynamics study. *Carbon*, 121:380–388.

[215] Marshall, M. S., Steele, R. P., Thanthiriwatte, K. S., and Sherrill, C. D. (2009). Potential energy curves for cation-π interactions: Off-axis configurations are also attractive. *Journal of Physical Chemistry A*, 113(48):13628–13632.

[216] Martin, J. W., Botero, M., Slavchov, R. I., Bowal, K., Akroyd, J., Mosbach, S., and Kraft, M. (2018). Flexoelectricity and the formation of carbon nanoparticles in flames. *Journal of Physical Chemistry C*.

[217] Martin, J. W., Bowal, K., Menon, A., Slavchov, R. I., Akroyd, J., Mosbach, S., and Kraft, M. (2019a). Polar curved polycyclic aromatic hydrocarbons in soot formation. *Proceedings of the Combustion Institute*, 37(1):1117–1123.

[218] Martin, J. W., de Tomas, C., Suarez-Martinez, I., Kraft, M., and Marks, N. A. (2019b). Topology of disordered 3D graphene networks. *Physical Review Letters*, 123(11):116105.

[219] Martin, J. W., Hou, D., Menon, A., Akroyd, J., You, X., and Kraft, M. (2019c). Reactivity of polycyclic aromatic hydrocarbon soot precursors: Implications of localised π-radicals on rim-based pentagonal rings. *The Journal of Physical Chemistry C*, 143:26673–26682.

[220] Martin, J. W., Nyadong, L., Ducati, C., Manley-Harris, M., Marshall, A. G., and Kraft, M. (2019d). Nanostructure of gasification charcoal (biochar). *Environmental science & technology*, 53(7):3538–3546.

[221] Martin, J. W., Slavchov, R. I., Yapp, E. K., Akroyd, J., Mosbach, S., and Kraft, M. (2017). The polarization of polycyclic aromatic hydrocarbons curved by pentagon incorporation: The role of the flexoelectric dipole. *Journal of Physical Chemistry C*, 121(48):27154–27163.

[222] Martinez, L., Andrade, R., Birgin, E. G., and Martinez, J. M. (2009). PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry*, 30(13):2157–2164.

[223] Menon, A., Dreyer, J. A., Martin, J. W., Akroyd, J., Robertson, J., and Kraft, M. (2019). Optical band gap of cross-linked, curved, and radical polyaromatic hydrocarbons. *Physical Chemistry Chemical Physics*, 21(29):16240–16251.

[224] Merchan-Merchan, W., Granados Sanmiguel, S., and McCollam, S. (2012). Analysis of soot particles derived from biodiesels and diesel fuel air-flames. *Fuel*, 102:525–535.

[225] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.

[226] Michelsen, H. (2017). Probing soot formation, chemical and physical evolution, and oxidation: A review of in situ diagnostic techniques and needs. *Proceedings of the Combustion Institute*, 36(1):717–735.

[227] Middleton, T. F., Hernández-Rojas, J., Mortenson, P. N., and Wales, D. J. (2001). Crystals of binary Lennard-Jones solids. *Physical Review B*, 64(18):184201.

[228] Migliorini, F., Thomson, K., and Smallwood, G. (2011). Investigation of optical properties of aging soot. *Applied Physics B*, 104(2):273–283.

[229] Miller, H., Vertes, A., Golden, R. J., and Korte, A. R. (2020). The molecular composition of soot. *Angewandte Chemie*, 132(11):4514–4520.

[230] Miller, J. H. (1991). The kinetics of polynuclear aromatic hydrocarbon agglomeration in flames. *Symposium (International) on Combustion*, 23:91–98.

[231] Miller, J. H., Herdman, J. D., Green, C. D., and Webster, E. M. (2013). Experimental and computational determinations of optical band gaps for pah and soot in a n2-diluted, ethylene/air non-premixed flame. *Proceedings of the Combustion Institute*, 34(2):3669–3675.

[232] Misquitta, A. J., Podeszwa, R., Jeziorski, B., and Szalewicz, K. (2005). Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional theory. *Journal of Chemical Physics*, 123:214103.

[233] Misquitta, A. J. and Stone, A. J. (2008). Dispersion energies for small organic molecules: first row atoms. *Molecular Physics*, 106:1631 – 1643.

[234] Misquitta, A. J. and Stone, A. J. (2016). Ab initio atom-atom potentials using CamCASP: Theory and application to many-body models for the pyridine dimer. *Journal of Chemical Theory and Computation*, 12(9):4184–4208. PMID: 27467814.

[235] Mitra, T., Zhang, T., Sediako, A. D., and Thomson, M. J. (2019). Understanding the formation and growth of polycyclic aromatic hydrocarbons (pahs) and young soot from n-dodecane in a sooting laminar coflow diffusion flame. *Combustion and Flame*, 202:33–42.

[236] Mravlak, M., Kister, T., Kraus, T., and Schilling, T. (2016). Structure diagram of binary Lennard-Jones clusters. *The Journal of Chemical Physics*, 145(2):024302.

[237] Müller, J.-O., Frank, B., Jentoft, R. E., Schlögl, R., and Su, D. S. (2012). The oxidation of soot particulate in the presence of NO_2. *Catal. Today*, 191(1):106–111.

[238] Myhre, G. (2009). Consistency between satellite-derived and modeled estimates of the direct aerosol effect. *Science*, 325(5937):187–190.

[239] Nagpal, P., Jamal, S., Singh, H., Ali, W., Tanweer, S., Sharma, R., Grover, A., and Grover, S. (2020). Long-range replica exchange molecular dynamics guided drug repurposing against tyrosine kinase ptka of mycobacterium tuberculosis. *Scientific Reports*, 10(1):1–11.

[240] Neves, A. R., Fernandes, P. A., and Ramos, M. J. (2011). The accuracy of density functional theory in the description of cation-π and π-hydrogen bond interactions. *Journal of Chemical Theory and Computation*, 7(7):2059–2067.

[241] Nishimoto, T., Takahashi, Y., Miyama, S., Furuta, T., and Sakurai, M. (2019). Replica exchange molecular dynamics simulation study on the mechanism of desiccation-induced structuralization of an intrinsically disordered peptide as a model of lea proteins. *Biophysics and physicobiology*, 16:196–204.

[242] Nosé, S. (1984). A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 52(2):255–268.

[243] Noya, E. G. and Doye, J. P. (2006). Structural transitions in the 309-atom magic number Lennard-Jones cluster. *The Journal of Chemical Physics*, 124(10):104503.

[244] Nymeyer, H. (2008). How efficient is replica exchange molecular dynamics? an analytic approach. *Journal of Chemical Theory and Computation*, 4(4):626–636.

[245] Oakley, M. T. and Johnston, R. L. (2014). Energy landscapes and global optimization of self-assembling cyclic peptides. *Journal of Chemical Theory and Computation*, 10(4):1810–1816.

[246] Obolensky, O. I., Semenikhina, V. V., Solov'yov, A. V., and Greiner, W. (2007). Interplay of electrostatic and van der Waals forces in coronene dimer. *International Journal of Quantum Chemistry*, 107(6):1335–1343.

[247] Öktem, B., Tolocka, M. P., Zhao, B., Wang, H., and Johnston, M. V. (2005). Chemical species associated with the early stage of soot growth in a laminar premixed ethylene–oxygen–argon flame. *Combustion and Flame*, 142(4):364–373.

[248] Olafusi, O. S., Sadiku, E. R., Snyman, J., Ndambuki, J. M., and Kupolati, W. K. (2019). Application of nanotechnology in concrete and supplementary cementitious materials: a review for sustainable construction. *SN Applied Sciences*, 1(6):580.

[249] Parent, P., Laffon, C., Marhaba, I., Ferry, D., Regier, T., Ortega, I., Chazallon, B., Carpentier, Y., and Focsa, C. (2016). Nanoscale characterization of aircraft soot: A high-resolution transmission electron microscopy, Raman spectroscopy, X-ray photoelectron and near-edge X-ray absorption spectroscopy study. *Carbon*, 101:86–100.

[250] Pascazio, L., Martin, J. W., Botero, M. L., Sirignano, M., D'Anna, A., and Kraft, M. (2019). Mechanical properties of soot particles: The impact of crosslinked polycyclic aromatic hydrocarbons. *Combustion Science and Technology*, pages 1–21.

[251] Pascazio, L., Martin, J. W., Bowal, K., Akroyd, J., and Kraft, M. (2020). Exploring the internal structure of soot particles using nanoindentation: A reactive molecular dynamics study. *Combustion and Flame*, 219:45–56.

[252] Pascazio, L., Sirignano, M., and Anna, A. D. (2017). Simulating the morphology of clusters of polycyclic aromatic hydrocarbons: The influence of the intermolecular potential. *Combustion and Flame*, 185:53–62.

[253] Patel, T. (1996). Stone age picassos. *New Scientist*, 151:32–35.

[254] Pedersen, A. and Jónsson, H. (2010). Distributed implementation of the adaptive kinetic Monte Carlo method. *Mathematics and Computers in Simulation*, 80(7):1487–1498.

[255] Petrukhina, M. A., Andreini, K. W., Mack, J., and Scott, L. T. (2005). X-ray quality geometries of geodesic polyarenes from theoretical calculations: What levels of theory are reliable? *Journal of Organic Chemistry*, 70(14):5713–5716.

[256] Petrukhina, M. A., Andreini, K. W., Peng, L., and Scott, L. T. (2004). Hemibuckminsterfullerene $C_{30}H_{12}$: X-ray crystal structures of the parent hydrocarbon and of the two-dimensional organometallic network $\{[Rh_2(O_2CCF_3)_4]_3\ (C_{30}H_{12})\}$. *Angewandte Chemie International Edition*, 43(41):5477–5481.

[257] Podeszwa, R. (2010). Interactions of graphene sheets deduced from properties of polycyclic aromatic hydrocarbons. *The Journal of Chemical Physics*, 132(4):044704.

[258] Podeszwa, R., Bukowski, R., and Szalewicz, K. (2006). Potential energy surface for the benzene dimer and perturbational analysis of π-π interactions. *Journal of Physical Chemistry A*, 110(34):10345–10354.

[259] Podeszwa, R. and Szalewicz, K. (2008). Physical origins of interactions in dimers of polycyclic aromatic hydrocarbons. *Physical Chemistry Chemical Physics*, 10:2735–2746.

[260] Pope, C. J. and Howard, J. B. (1997). Simultaneous particle and molecule modeling (spamm): An approach for combining sectional aerosol equations and elementary gas-phase reactions. *Aerosol science and technology*, 27(1):73–94.

[261] Pope III, C. A. and Dockery, D. W. (2006). Health effects of fine particulate air pollution: lines that connect. *Journal of the air & waste management association*, 56(6):709–742.

[262] Porter, G. (1953). Carbon formation in the combustion wave. *Symposium (International) on Combustion*, 4(1):248–252.

[263] Rahman, A. (1964). Correlations in the motion of atoms in liquid argon. *Physical review*, 136(2A):A405.

[264] Raj, A., Man, P. L., Totton, T. S., Sander, M., Shirley, R. A., and Kraft, M. (2010a). New polycyclic aromatic hydrocarbon (pah) surface processes to improve the model prediction of the composition of combustion-generated pahs and soot. *Carbon*, 48(2):319–332.

[265] Raj, A., Sander, M., Janardhanan, V., and Kraft, M. (2010b). A study on the coagulation of polycyclic aromatic hydrocarbon clusters to determine their collision efficiency. *Combustion and Flame*, 157(3):523–534.

[266] Rapacioli, M., Calvo, F., Joblin, C., Parneix, P., Toublanc, D., and Spiegelman, F. (2006). Formation and destruction of polycyclic aromatic hydrocarbon clusters in the interstellar medium. *Astronomy & Astrophysics*, 460(2):519–531.

[267] Rapacioli, M., Calvo, F., Spiegelman, F., Joblin, C., and Wales, D. J. (2005). Stacked clusters of polycyclic aromatic hydrocarbon molecules. *The Journal of Physical Chemistry A*, 109(11):2487–2497.

[268] Rathore, N., Chopra, M., and De Pablo, J. J. (2005). Optimal allocation of replicas in parallel tempering simulations. *The Journal of Chemical Physics*, 122(10):24111–174903.

[269] Rice, A. M., Dolgopolova, E. A., Yarbrough, B. J., Leith, G. A., Martin, C. R., Stephenson, K. S., Heugh, R. A., Brandt, A. J., Chen, D. A., Karakalos, S. G., et al. (2018). Stack the bowls: Tailoring the electronic structure of corannulene-integrated crystalline materials. *Angewandte Chemie International Edition*, 57(35):11310–11315.

[270] Richter, H. and Howard, J. B. (2000). Formation of polycyclic aromatic hydrocarbons and their growth to soot—a review of chemical reaction pathways. *Progress in Energy and Combustion science*, 26(4-6):565–608.

[271] Rissler, J., Swietlicki, E., Bengtsson, A., Boman, C., Pagels, J., Sandström, T., Blomberg, A., and Löndahl, J. (2012). Experimental determination of deposition of diesel exhaust particles in the human respiratory tract. *Journal of Aerosol Science*, 48:18–33.

[272] Roch, L. M., Zoppi, L., Siegel, J. S., and Baldridge, K. K. (2017). Indenocorannulene-based materials: Effect of solid-state packing and intermolecular interactions on optoelectronic properties. *The Journal of Physical Chemistry C*, 121(2):1220–1234.

[273] Rogan, J., Varas, A., Valdivia, J. A., and Kiwi, M. (2013). A strategy to find minimal energy nanocluster structures. *Journal of Computational Chemistry*, 34(29):2548–2556.

[274] Rondina, G. G. and Da Silva, J. L. (2013). Revised basin-hopping Monte Carlo algorithm for structure optimization of clusters and nanoparticles. *Journal of Chemical Information and Modeling*, 53(9):2282–2298.

[275] Rossi, G. and Ferrando, R. (2006). Global optimization by excitable walkers. *Chemical Physics Letters*, 423(1-3):17–22.

[276] Rossi, G., Rapallo, A., Mottet, C., Fortunelli, A., Baletto, F., and Ferrando, R. (2004). Magic polyicosahedral core-shell clusters. *Physical Review Letters*, 93(10):105503.

[277] Russo, C., Tregrossi, A., and Ciajolo, A. (2015). Dehydrogenation and growth of soot in premixed flames. *Proceedings of the Combustion Institute*, 35(2):1803–1809.

[278] Ruuska, H. and Pakkanen, T. A. (2001). Ab initio study of interlayer interaction of graphite: Benzene-coronene and coronene dimer two-layer models. *The Journal of Chemical Physics B*, 105(39):9541–9547.

[279] Sabbah, H., Biennier, L., Klippenstein, S. J., Sims, I. R., and Rowe, B. R. (2010). Exploring the role of pahs in the formation of soot: Pyrene dimerization. *The Journal of Physical Chemistry Letters*, 1(19):2962–2967.

[280] Safaei, S., Hendy, S. C., and Willmott, G. R. (2020). Stability of amphiphilic janus dimers in shear flow: a molecular dynamics study. *Soft Matter*, 16(30):7116–7125.

[281] Saffaripour, M., Veshkini, A., Kholghy, M., and Thomson, M. J. (2014). Experimental investigation and detailed modeling of soot aggregate formation and size distribution in laminar coflow diffusion flames of jet a-1, a synthetic kerosene, and n-decane. *Combustion and Flame*, 161(3):848–863.

[282] Sakurai, H., Daiko, T., Sakane, H., Amaya, T., and Hirao, T. (2005). Structural elucidation of sumanene and generation of its benzylic anions. *Journal of the American Chemical Society*, 127(33):11580–11581.

[283] Saldinger, J. C., Wang, Q., Elvati, P., and Violi, A. (2020). Characterizing the diversity of aromatics in a coflow diffusion jet a-1 surrogate flame. *Fuel*, 268:117198.

[284] Sanner, M. F., Olson, A. J., and Spehner, J.-C. (1996). Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers*, 38(3):305–320.

[285] Sanyal, S., Manna, A. K., and Pati, S. K. (2014). Functional corannulene: diverse structures, enhanced charge transport, and tunable optoelectronic properties. *ChemPhysChem*, 15(5):885–893.

[286] Scanlon, L., Balbuena, P., Zhang, Y., Sandi, G., Back, C., Feld, W., Mack, J., Rottmayer, M., and Riepenhoff, J. (2006). Investigation of corannulene for molecular hydrogen storage via computational chemistry and experimentation. *The Journal of Physical Chemistry B*, 110(15):7688–7694.

[287] Schenk, M., Lieb, S., Vieker, H., Beyer, A., Gölzhäuser, A., Wang, H., and Kohse-Höinghaus, K. (2015). Morphology of nascent soot in ethylene flames. *Proceedings of the Combustion Institute*, 35(2):1879–1886.

[288] Schröder, D., Loos, J., Schwarz, H., Thissen, R., Preda, D. V., Scott, L. T., Caraiman, D., Frach, M. V., and Böhme, D. K. (2001). Single and double ionization of corannulene and coronene. *Helvetica Chimica Acta*, 84(6):1625–1634.

[289] Schueler-Furman, O., Wang, C., Bradley, P., Misura, K., and Baker, D. (2005). Progress in modeling of protein structures and interactions. *Science*, 310(5748):638–642.

[290] Schuetz, C. A. and Frenklach, M. (2002). Nucleation of soot: molecular dynamics simulations of pyrene dimerization. *Proceedings of the Combustion Institute*, 29(2):2307–2314.

[291] Schulz, F., Commodo, M., Kaiser, K., De Falco, G., Minutolo, P., Meyer, G., Andrea, D., and Gross, L. (2019). Insights into incipient soot formation by atomic force microscopy. *Proceedings of the Combustion Institute*, 37(1):885–892.

[292] Schutte, W., Tielens, A., and Allamandola, L. (1993). Theoretical modeling of the infrared fluorescence from interstellar polycyclic aromatic hydrocarbons. *The Astrophysical Journal*, 415:397–414.

[293] Schyman, P. and Jorgensen, W. L. (2013). Exploring adsorption of water and ions on carbon surfaces using a polarizable force field. *Journal of Physical Chemistry Letters*, 4(3):468–474.

[294] Scott, L. T., Bronstein, H. E., Preda, D. V., Ansems, R. B., Bratcher, M. S., and Hagen, S. (1999). Geodesic polyarenes with exposed concave surfaces. *Pure and Applied Chemistry*, 71(2):209–219.

[295] Scott, L. T., Hashemi, M. M., and Bratcher, M. S. (1992). Corannulene bowl-to-bowl inversion is rapid at room temperature. *American Chemical Society*, pages 1920–1921.

[296] Senftle, T. P., Hong, S., Islam, M. M., Kylasa, S. B., Zheng, Y., Shin, Y. K., Junkermeier, C., Engel-Herbert, R., Janik, M. J., Aktulga, H. M., Verstraelen, T., Grama, A., and van Duin, A. C. T. (2016). The ReaxFF reactive force-field: development, applications and future directions. *npj Computational Materials*, 2:1–14.

[297] Sgro, L., De Filippo, A., Lanzuolo, G., and D'Alessio, A. (2007). Characterization of nanoparticles of organic carbon (noc) produced in rich premixed flames by differential mobility analysis. *Proceedings of the Combustion Institute*, 31(1):631–638.

[298] Shostak, S. L., Ebenstein, W. L., and Muenter, J. S. (1991). The dipole moment of water. I. Dipole moments and hyperfine properties of H_2O and HDO in the ground and excited vibrational states. *The Journal of Chemical Physics*, 94(9):5875.

[299] Shu, Q., Yang, Y., Zhai, Y., Sun, D. Y., Xiang, H. J., and Gong, X. G. (2012). Size-dependent melting behavior of iron nanoparticles by replica exchange molecular dynamics. *Nanoscale*, 4(20):6307–6311.

[300] Simonsson, J., Olofsson, N. E., Bladh, H., Sanati, M., and Bengtsson, P. E. (2017). Influence of potassium and iron chloride on the early stages of soot formation studied using imaging LII/ELS and TEM techniques. *Proceedings of the Combustion Institute*, 36(1):853–860.

[301] Sindhikara, D., Meng, Y., and Roitberg, A. E. (2008). Exchange frequency in replica exchange molecular dynamics. *The Journal of Chemical Physics*, 128(2):01B609.

[302] Sindhikara, D. J., Emerson, D. J., and Roitberg, A. E. (2010). Exchange often and properly in replica exchange molecular dynamics. *Journal of Chemical Theory and Computation*, 6(9):2804–2808.

[303] Singh, J., Balthasar, M., Kraft, M., and Wagner, W. (2005). Stochastic modeling of soot particle size and age distributions in laminar premixed flames. *Proceedings of the Combustion Institute*, 30(1):1457–1465.

[304] Sinnokrot, M. O. and Sherrill, C. D. (2004). Highly accurate coupled cluster potential energy curves for the benzene dimer: Sandwich, T-shaped, and parallel-displaced configurations. *Journal of Physical Chemistry A*, 108(46):10200–10207.

[305] Smeeton, L. C., Farrell, J. D., Oakley, M. T., Wales, D. J., and Johnston, R. L. (2015). Structures and energy landscapes of hydrated sulfate clusters. *Journal of Chemical Theory and Computation*, 11(5):2377–2384.

[306] Srilomsak, M. and Hanamura, K. (2020). Time-lapse visualization of diesel particulate filter active-regeneration using field emission scanning electron microscopy. *Journal of Microscopy*.

[307] Stein, S. E. and Fahr, A. (1985). High-temperature stabilities of hydrocarbons. *The Journal of Physical Chemistry*, 89(17):3714–3725.

[308] Stone, A. (2013). *The theory of intermolecular forces*. OUP Oxford.

[309] Stone, A. J. (2005). Distributed multipole analysis: Stability for large basis sets. *Journal of Chemical Theory and Computation*, 1:1128–1132.

[310] Stone, A. J., Alderton, &. M., and Alderton, M. (1985). Distributed multipole analysis: Methods and applications. *Molecular Physics*, 56(5):1047–1064.

[311] Stone, A. J., Dullweber, A., Engkvist, O., Fraschini, E., Hodges, M. P., Meredith, A. W., Popelier, P. L. A., and Wales, D. J. (2017). ORIENT: a program for studying interactions between molecules, version 4.9.

[312] Strodel, B., Lee, J. W., Whittleston, C. S., and Wales, D. J. (2010). Transmembrane structures for alzheimer's aβ_{1-42} oligomers. *Journal of the American Chemical Society*, 132(38):13300–13312.

[313] Sugita, Y. and Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, 314:141–151.

[314] Swendsen, R. and Wang, J. (1986). Replica Monte Carlo simulation of spin glasses. *Phys. Rev. Lett.*, 57(21):2607–2609.

[315] Sygula, A., Folsom, H. E., Sygula, R., Abdourazak, A. H., Marcinow, Z., Fronczek, F. R., and Rabideau, P. W. (1994). Bowl stacking in curved polynuclear aromatic hydrocarbons: crystal and molecular structure of cyclopentacorannulene. *Journal of the Chemical Society, Chemical Communications*, 22:2571–2572.

[316] Sygula, A. and Saebø, S. (2009). π-π stacking of curved carbon networks: The corannulene dimer. *International Journal of Quantum Chemistry*, 109(1):65–72.

[317] Takeuchi, H. (2013). Structures, stability, and growth sequence patterns of small homoclusters of naphthalene, anthracene, phenanthrene, phenalene, naphthacene, and pyrene. *Computational and Theoretical Chemistry*, 1021:84–90.

[318] Tang, K. T. and Toennies, J. P. (1984). An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients. *The Journal of Chemical Physics*, 80:3726–3741.

[319] Tang, Q., Cai, R., You, X., and Jiang, J. (2017). Nascent soot particle size distributions down to 1 nm from a laminar premixed burner-stabilized stagnation ethylene flame. *Proceedings of the Combustion Institute*, 36(1):993–1000.

[320] Teini, P. D., Karwat, D. M. A., and Atreya, A. (2011). Observations of nascent soot: Molecular deposition and particle morphology. *Combustion and Flame*, 158:2045–2055.

[321] Terzyk, A. P., Furmaniak, S., Gauden, P. A., Harris, P. J., and Kowalczyk, P. (2012). *Virtual Porous Carbons*. Elsevier.

[322] Tielens, A. G. (2008). Interstellar polycyclic aromatic hydrocarbon molecules. *Annual Review of Astronomy and Astrophysics*, 46:289–337.

[323] Totton, T. S., Chakrabarti, D., Misquitta, A. J., Sander, M., Wales, D. J., and Kraft, M. (2010a). Modelling the internal structure of nascent soot particles. *Combustion and Flame*, 157(5):909–914.

[324] Totton, T. S., Misquitta, A. J., and Kraft, M. (2010b). A first principles development of a general anisotropic potential for polycyclic aromatic hydrocarbons. *Journal of Chemical Theory and Computation*, 6(3):683–695.

[325] Totton, T. S., Misquitta, A. J., and Kraft, M. (2011a). A transferable electrostatic model for intermolecular interactions between polycyclic aromatic hydrocarbons. *Chemical Physics Letters*, 510(1-3):154–160.

[326] Totton, T. S., Misquitta, A. J., and Kraft, M. (2011b). Assessing the polycyclic aromatic hydrocarbon anisotropic potential with application to the exfoliation energy of graphite. *The Journal of Physical Chemistry A*, 115(46):13684–13693.

[327] Totton, T. S., Misquitta, A. J., and Kraft, M. (2012). A quantitative study of the clustering of polycyclic aromatic hydrocarbons at high temperatures. *Physical Chemistry Chemical Physics*, 14(12):4081–4094.

[328] Tsai, C. and Jordan, K. (1993). Use of an eigenmode method to locate the stationary points on the potential energy surfaces of selected argon and water clusters. *The Journal of Physical Chemistry*, 97(43):11227–11237.

[329] Tsuzuki, S., Honda, K., Uchimaru, T., and Mikami, M. (2004). High-level ab initio computations of structures and interaction energies of naphthalene dimers: Origin of attraction and its directionality. *Journal of Chemical Physics*, 120(2):647–659.

[330] Tsuzuki, S., Honda, K., Uchimaru, T., Mikami, M., and Tanabe, K. (2002). Origin of attraction and directionality of the π/π interaction: Model chemistry calculations of benzene dimer interaction. *Journal of the American Chemical Society*, 124(1):104–112.

[331] Tsuzuki, S., Yoshida, M., Uchimaru, T., and Mikami, M. (2001). The origin of the cation/π interaction: The significant importance of the induction in Li^+ and Na^+ complexes. *The Journal of Physical Chemistry A*, 105(4):769–773.

[332] Ugarte, D. (1992). Curling and closure of graphitic networks under electron-beam irradiation. *Nature*, 359(6397):707–709.

[333] Ugarte, D. (1995). Onion-like graphitic particles. *Carbon*, 33(7):989–993.

[334] Van de Waal, B. W. (1983). Calculated ground-state structures of 13-molecule clusters of carbon dioxide, methane, benzene, cyclohexane, and naphthalene. *The Journal of Chemical Physics*, 79(8):3948–3961.

[335] Vander Wal, R. L. and Tomasek, A. J. (2003). Soot oxidation: dependence upon initial nanostructure. *Combustion and Flame*, 134(1-2):1–9.

[336] Verlet, L. (1967). Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical Review*, 159(1):98–103.

[337] Verma, P., Pickering, E., Savic, N., Zare, A., Brown, R., and Ristovski, Z. (2019). Comparison of manual and automatic approaches for characterisation of morphology and nanostructure of soot particles. *Journal of Aerosol Science*, 136:91–105.

[338] Violi, A. (2004). Modeling of soot particle inception in aromatic and aliphatic premixed flames. *Combustion and Flame*, 139(4):279–287.

[339] Violi, A., Kubota, A., Truong, T., Pitz, W., Westbrook, C., and Sarofim, A. (2002). A fully integrated kinetic monte carlo/molecular dynamics approach for the simulation of soot precursor growth. *Proceedings of the Combustion Institute*, 29(2):2343–2349.

[340] Wainer, H. and Friendly, M. (2018). Visual revelations: ancient visualizations. *Chance*, 31(2):62–64.

[341] Wales, D. J. (2011). Energy landscapes and structure prediction using basin-hopping. In Oganov, A. R., editor, *Modern Methods of Crystal Structure Prediction*, chapter 2, pages 29–54. Wiley-VCH.

[342] Wales, D. J. (2013). Surveying a complex potential energy landscape: Overcoming broken ergodicity using basin-sampling. *Chemical Physics Letters*, 584:1–9.

[343] Wales, D. J. (2018). Exploring energy landscapes. *Annual Review of Physical Chemistry*, 69:401–425.

[344] Wales, D. J. and Doye, J. P. K. (1997). Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116.

[345] Wales, D. J., Miller, M. A., and Walsh, T. R. (1998). Archetypal energy landscapes. *Nature*, 394(6695):758.

[346] Wales, D. J. and Scheraga, H. A. (1999). Global optimization of clusters, crystals, and biomolecules. *Science*, 285(5432):1368–1372.

[347] Wan, K., Chen, D., and Wang, H. (2018). On imaging nascent soot by transmission electron microscopy. *Combustion and Flame*, 198:260–266.

[348] Wang, A., Lou, H. H., Chen, D., Yu, A., Dang, W., Li, X., Martin, C., Damodara, V., and Patki, A. (2016). Combustion mechanism development and cfd simulation for the prediction of soot emission during flaring. *Frontiers of Chemical Science and Engineering*, 10(4):459–471.

[349] Wang, B.-T., Petrukhina, M. A., and Margine, E. R. (2015). Electronic transport properties of selected carbon π-bowls with different size, curvature and solid state packing. *Carbon*, 94:174–180.

[350] Wang, C., Huddle, T., Huang, C.-H., Zhu, W., Vander Wal, R. L., Lester, E. H., and Mathews, J. P. (2017). Improved quantification of curvature in high-resolution transmission electron microscopy lattice fringe micrographs of soots. *Carbon*, 117:174–181.

[351] Wang, C. S., Bartelt, N. C., Ragan, R., and Thürmer, K. (2018a). Revealing the molecular structure of soot precursors. *Carbon*, 129:537–542.

[352] Wang, H. (2011). Formation of nascent soot and other condensed-phase materials in flames. *Proceedings of the Combustion Institute*, 33(1):41 – 67.

[353] Wang, H., Xu, R., Wang, K., Bowman, C. T., Hanson, R. K., Davidson, D. F., Brezinsky, K., and Egolfopoulos, F. N. (2018b). A physics-based approach to modeling real-fuel combustion chemistry-i. evidence from experiments, and thermodynamic, chemical kinetic and statistical considerations. *Combustion and Flame*, 193:502–519.

[354] Wang, Q.-D., Wang, J.-B., Li, J.-Q., Tan, N.-X., and Li, X.-Y. (2011a). Reactive molecular dynamics simulation and chemical kinetic modeling of pyrolysis and combustion of n-dodecane. *Combustion and Flame*, 158:217–226.

[355] Wang, W., Liang, S., Yu, T., Li, D., Li, Y., and Han, X. (2011b). The study of interaction between graphene and metals by raman spectroscopy. *Journal of Applied Physics*, 109(7):07C501.

[356] Wen, J. Z., Thomson, M., Park, S., Rogak, S., and Lightstone, M. (2005). Study of soot growth in a plug flow reactor using a moving sectional model. *Proceedings of the Combustion Institute*, 30(1):1477–1484.

[357] Wentzel, M., Gorzawski, H., Naumann, K.-H., Saathoff, H., and Weinbruch, S. (2003). Transmission electron microscopical and aerosol dynamical characterization of soot aerosols. *Journal of aerosol science*, 34(10):1347–1370.

[358] Wersborg, B. L., Howard, J. B., and Williams, G. C. (1973). Physical mechanisms in carbon formation in flames. *Symposium (International) on Combustion*, 14(1):929–940.

[359] Whitesides, R. and Frenklach, M. (2010). Detailed kinetic monte carlo simulations of graphene-edge growth. *The Journal of Physical Chemistry A*, 114(2):689–703.

[360] Williams, D. E. (1999a). Improved intermolecular force field for crystalline hydrocarbons containing four- or three-coordinated carbon. *Journal of Molecular Structure*, 485:321–347.

[361] Williams, D. E. (1999b). Improved intermolecular force field for crystalline hydrocarbons containing four- or three-coordinated carbon. *Journal of Molecular Structure*, 485-486:321–347.

[362] Williams, D. E. (2001). Improved intermolecular force field for molecules containing H, C, N, and O atoms, with applications to nucleoside and peptide crystals. *Journal of Computational Chemistry*, 22:1154–1166.

[363] Wu, Y.-T. and Siegel, J. S. (2006). Aromatic molecular-bowl hydrocarbons: synthetic derivatives, their structures, and physical properties. *Chemical reviews*, 106(12):4843–4867.

[364] Xu, R., Wang, K., Banerjee, S., Shao, J., Parise, T., Zhu, Y., Wang, S., Movaghar, A., Lee, D. J., Zhao, R., et al. (2018). A physics-based approach to modeling real-fuel combustion chemistry–ii. reaction kinetic models of jet and rocket fuels. *Combustion and Flame*, 193:520–537.

[365] Yehliu, K., Vander Wal, R. L., Armas, O., and Boehman, A. L. (2012). Impact of fuel formulation on the nanostructure and reactivity of diesel soot. *Combustion and Flame*, 159(12):3597–3606.

[366] Yon, J., Ouf, F.-X., Hebert, D., Mitchell, J. B., Teuscher, N., Le Garrec, J.-L., Bescond, A., Baumann, W., Ourdani, D., Bizien, T., et al. (2018). Investigation of soot oxidation by coupling lii, saxs and scattering measurements. *Combustion and Flame*, 190:441–453.

[367] Zabula, A. V., Spisak, S. N., Filatov, A. S., Rogachev, A. Y., and Petrukhina, M. A. (2018). Record alkali metal intercalation by highly charged corannulene. *Accounts of Chemical Research*, 51(6):1541–1549.

[368] Zhang, C., Zhang, C., Ma, Y., and Xue, X. (2015). Imaging the C black formation by acetylene pyrolysis with molecular reactive force field simulations. *Physical Chemistry Chemical Physics*, 17(17):11469–11480.

[369] Zhang, J., Grzybowski, B. A., and Granick, S. (2017). Janus particle synthesis, assembly, and application. *Langmuir*, 33(28):6964–6977.

[370] Zhang, Y., Zhuo, J., Wu, Y., and Yao, Q. (2020). Molecular simulation of the adsorption behaviors of CO_2/CH_4 in curvature, planar, and mixture models. *Energy & Fuels*, 34(4):4153–4161.

[371] Zhang, Z., Zhang, J., Chen, N., and Qu, L. (2012). Graphene quantum dots: an emerging material for energy-related applications and beyond. *Energy & Environmental Science*, 5(10):8869–8890.

[372] Zhao, L., Yang, T., Kaiser, R. I., Troy, T. P., Ahmed, M., Belisario-Lara, D., Ribeiro, J. M., and Mebel, A. M. (2017). Combined experimental and computational study on the unimolecular decomposition of jp-8 jet fuel surrogates. i. n-decane (n-c10h22). *The Journal of Physical Chemistry A*, 121(6):1261–1280.

[373] Zhao, Y. and Truhlar, D. G. (2008). A prototype for graphene material simulation: Structures and interaction potentials of coronene dimers. *The Journal of Physical Chemistry C*, 112(11):4061–4067.

[374] Zhong, Q., Mao, Q., Zhang, L., Xiang, J., Xiao, J., and Mathews, J. P. (2018). Structural features of Qingdao petroleum coke from HRTEM lattice fringes: Distributions of length, orientation, stacking, curvature, and a large-scale image-guided 3D atomistic representation. *Carbon*, 129:790–802.

[375] Zhou, L., Xiong, G., Zhang, M., Chen, L., Ding, S., and De Goey, L. (2017). Experimental study of polycyclic aromatic hydrocarbons (pahs) in n-heptane laminar diffusion flames from 1. 0 to 3.0 bar. *Fuel*, 209:265–273.

Appendix A

Scope of collaboration

The work in this thesis received the following contributions from co-authors.

Chapter 4: J. W. Martin provided helpful discussions on the methodology and feedback on manuscript drafts.

Chapter 5: D. Chen provided scripts to run solvent-excluded surface analysis of PAH clusters and feedback on manuscript drafts. H. Wang ran the first solvent-excluded surface analysis calculations. L. Pascazio provided feedback on the manuscript draft.

Chapter 6: P. Grančič proposed the first version of the SEMC method and produced the first scripts to execute it. J. W. Martin provided feedback on the manuscript drafts.

Chapter 7: J. W. Martin provided feedback on the manuscript draft, A. J. Misquitta provided helpful discussions related to force field development and feedback on the manuscript drafts.

Chapter 8: J. W. Martin provided assistance with the density functional theory calculations and gave feedback on the manuscript drafts.

Appendix B

Supplementary Information for Chapter 7

B.1 Corannulene geometry and charges

Corannulene elements, geometry (x, y, z in nm) and charges, in the curPAHIP potential. The element X represents the virtual mass-less atoms used to describe the flexoelectric effect, which are held above and parallel to the pentagon ring using intramolecular forces.

```
35
CORANNULENE with curPAHIP charges
C          0.103   0.064   0.085    0.020
C          0.093  -0.078   0.085    0.020
C         -0.045  -0.112   0.085    0.020
C         -0.121   0.008   0.085    0.020
C         -0.030   0.117   0.085    0.020
C          0.205   0.129   0.022    0.177
C          0.310   0.049  -0.032   -0.201
C          0.300  -0.090  -0.032   -0.201
C          0.187  -0.156   0.022    0.177
C          0.143  -0.279  -0.032   -0.201
C          0.007  -0.313  -0.033   -0.201
C         -0.090  -0.225   0.022    0.177
C         -0.242   0.016   0.022    0.177
```

```
C       -0.296  -0.103  -0.032  -0.201
C       -0.222  -0.222  -0.032  -0.201
C       -0.280   0.142  -0.031  -0.201
C       -0.190   0.250  -0.031  -0.201
C       -0.059   0.235   0.023   0.177
C        0.049   0.310  -0.032  -0.201
C        0.178   0.258  -0.032  -0.201
H       -0.221   0.345  -0.072   0.133
H        0.031   0.408  -0.072   0.133
H        0.259   0.317  -0.073   0.133
H        0.398   0.097  -0.072   0.133
H        0.380  -0.149  -0.075   0.133
H        0.215  -0.348  -0.073   0.133
H       -0.024  -0.408  -0.074   0.133
H       -0.265  -0.312  -0.073   0.133
H       -0.396  -0.103  -0.074   0.133
H       -0.378   0.155  -0.073   0.133
X        0.103   0.064   0.132  -0.063
X        0.093  -0.078   0.132  -0.063
X       -0.045  -0.113   0.132  -0.063
X       -0.121   0.008   0.133  -0.063
X       -0.030   0.117   0.133  -0.063
```

B.2 curPAHIP parameters in SI units

Table B.1 Parameters of curPAHIP in SI units

Atom pair	ρ (nm)	α (nm^{-1})	C_6 (kJmol^{-1}nm^6)
C C	0.2993	35.49	0.0017458
C H	0.2610	33.18	0.00072665
H H	0.2180	26.54	0.00030083

B.3 K^+ interaction parameters in SI units

$\sigma = 0.51$ nm and $\varepsilon = 0.00009$ kJ/mol

Appendix C

Supplementary Information for Chapter 8

C.1 $C_{42}H_{14}$ molecule description

The following elements, geometry (x, y, z in nm) and charges are used for the $C_{42}H_{14}$ molecule (B) in the curPAHIP potential. The element X represents the virtual mass-less atoms used to describe the flexoelectric effect, which are held above and parallel to the pentagon ring using intramolecular forces.

```
66
2PENT15RING MOLECULE with curPAHIP charges
C    -0.200   0.000   0.135   -0.309
C    -0.317   0.000   0.043    0.425
C    -0.365  -0.132  -0.016   -0.126
C    -0.126  -0.124   0.159    0.041
C    -0.126   0.124   0.158    0.041
C     0.014   0.124   0.176    0.070
C     0.093   0.000   0.174   -0.190
C     0.014  -0.125   0.177    0.070
C     0.065  -0.244   0.122   -0.117
C    -0.486  -0.127  -0.098   -0.225
C    -0.543   0.000  -0.139    0.010
C    -0.486   0.127  -0.096   -0.225
C    -0.365   0.131  -0.016   -0.126
```

C	-0.277	0.252	-0.001	0.253
C	-0.162	0.242	0.091	-0.109
C	-0.044	0.318	0.073	0.039
C	0.065	0.244	0.122	-0.117
C	0.197	0.251	0.059	0.008
C	0.282	0.128	0.056	0.150
C	0.229	0.000	0.111	0.038
C	-0.029	0.421	-0.031	0.157
C	0.110	0.463	-0.060	-0.365
C	0.219	0.365	-0.036	0.346
C	-0.154	0.458	-0.100	-0.163
C	-0.275	0.373	-0.088	-0.210
C	0.284	-0.130	0.058	0.150
C	0.198	-0.252	0.060	0.008
C	0.401	0.126	-0.034	-0.116
C	0.347	0.366	-0.109	-0.392
C	0.435	0.249	-0.107	-0.017
C	0.220	-0.365	-0.036	0.346
H	-0.526	0.219	-0.137	0.128
H	-0.629	0.001	-0.205	0.090
H	0.369	0.449	-0.175	0.162
H	0.127	0.540	-0.134	0.161
H	-0.153	0.535	-0.176	0.116
H	-0.356	0.391	-0.157	0.133
C	0.402	-0.128	-0.032	-0.116
C	0.473	0.000	-0.054	-0.115
C	0.352	-0.369	-0.100	-0.392
C	0.437	-0.250	-0.105	-0.017
C	-0.278	-0.253	0.001	0.253
C	-0.162	-0.243	0.092	-0.109
C	-0.044	-0.318	0.072	0.039
C	-0.030	-0.418	-0.035	0.157
C	0.109	-0.458	-0.067	-0.365
C	-0.280	-0.378	-0.079	-0.210
C	-0.156	-0.455	-0.103	-0.163
H	-0.363	-0.398	-0.146	0.133

```
H    -0.153  -0.525  -0.186   0.116
H    -0.527  -0.218  -0.140   0.128
H     0.126  -0.534  -0.142   0.161
H     0.376  -0.453  -0.165   0.162
H     0.520  -0.246  -0.173   0.104
H     0.558  -0.001  -0.121   0.129
H     0.523   0.248  -0.171   0.104
X    -0.132   0.148   0.204   0.044
X    -0.170   0.266   0.137  -0.107
X    -0.051   0.340   0.119   0.041
X     0.060   0.267   0.169  -0.115
X     0.008   0.148   0.223   0.072
X    -0.050  -0.341   0.118   0.041
X    -0.169  -0.267   0.138  -0.107
X    -0.132  -0.148   0.205   0.044
X     0.060  -0.267   0.168  -0.115
X     0.009  -0.148   0.223   0.072
```

C.2 Cut-off distance sensitivities

The selection of the cut-off distance, R, influences the calculated average intermolecular distances, coordination numbers and alignment angles. Due to the disordered molecular arrangements of the homogeneous A clusters (that is, sandwich-type stacking is not present), these results are the most sensitive to the selection of R.

Figure C.1 shows the alignment angle distributions for all homogeneous clusters using four different cut-off distances. The influence of the cut-off distance is minimal in the B clusters, which show a very high proportion of molecules with at least one near neighbour at all cut-off distances (indicated by the inset percent values listed). At the higher cut-off distances, a second peak corresponding to further molecule layers (*i.e.* not the nearest neighbours alone) appears. The A clusters show some differences between angle distributions across the cut-off distances and the percent of molecules with a near neighbour increases dramatically between $R = 0.5$ nm and $R = 0.6$ nm. Note that for all systems, no neighbouring molecules are found using a cut-off distance of 0.4 nm or smaller.

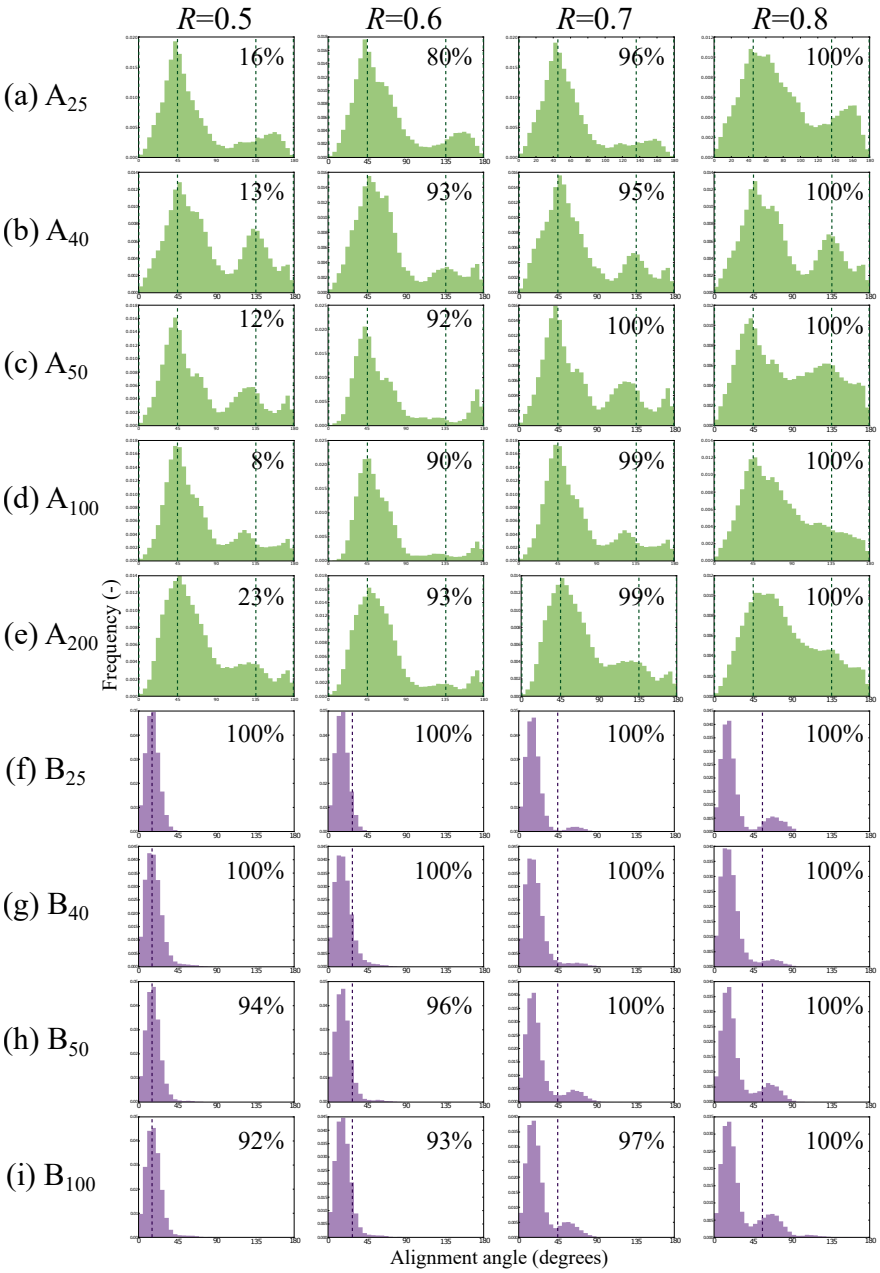


Fig. C.1 Alignment angle distributions at different cut-off distances, R, in nm for the following clusters: (a) A_{25}, (b) A_{40}, (c) A_{50}, (d) A_{100}, (e) A_{200}, (f) B_{25}, (g) B_{40}, (h) B_{50}, (i) B_{100}. Dashed lines correspond to the A crystal structure (for (a)-(e)) and the minimised B dimer (for (f)-(i)). Percent values in the upper right hand corners of each angle distribution refer to the percent of molecules within the cluster that have at least one near neighbour.

Table C.1 presents the average intermolecular distances of homogeneous clusters as a function of the cut-off distance used in analysis. For the A clusters, the cut-off distance selected influences the average intermolecular distance calculated. Intuitively, an increase in the

cut-off distance produces an increase in the average intermolecular distance, since the cut-off distance simply increases the range between 'neighbouring' molecules. At low cut-off values ($R = 0.5$ nm and $R = 0.6$ nm) very few molecule pairs exist and so these averages are not a clear picture of the true average intermolecular spacing throughout the cluster. At $R = 0.7$ nm, the majority of molecules ($\geq 95\%$) have at least one near neighbour and therefore this provides the best indication of the cluster average. In contrast, the cut-off distance does not have a large impact on the average intermolecular spacing of homogeneous B molecules. This highlights the highly stacked configuration of these molecules.

Table C.1 Average molecular intermolecular distance, in nm, for homogeneous clusters with varying cut-off values, R in nm, used.

Cluster	Intermolecular distance (nm)			
	$R = 0.5$	$R = 0.6$	$R = 0.7$	$R = 0.8$
A_{25}	0.45	0.55	0.59	0.68
A_{40}	0.45	0.55	0.59	0.67
A_{50}	0.47	0.55	0.60	0.67
A_{100}	0.48	0.55	0.59	0.68
A_{200}	0.47	0.55	0.60	0.68
B_{25}	0.44	0.44	0.45	0.49
B_{40}	0.45	0.45	0.46	0.47
B_{50}	0.45	0.45	0.48	0.50
B_{100}	0.45	0.45	0.48	0.52

C.3 Corannulene crystal structure

Several of the structural metrics used in this paper were applied to the known crystal structure of corannulene in order to provide a benchmark. As early as 1975, it was shown through X-ray analysis that corannulene crystallises in a close-packed structure mostly stabilised through CH-π interactions, in space group $P2_1/c$ [139] (see Fig C.2).

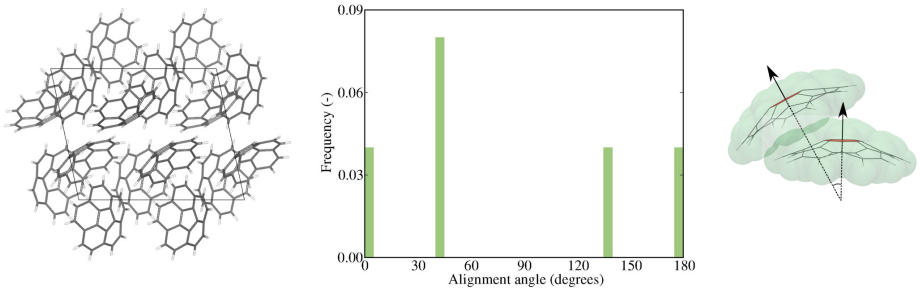


Fig. C.2 Snapshot of corannulene crystal structure (left) with alignment angle distribution (centre) and a schematic illustrating an alignment angle of 30° between two neighbouring A.

A crystal structure containing 18 molecules from X-ray data provided by Petrukhina et al. [255] through the Cambridge Crystallographic Data Centre provides an independent verification of the analyses used in Chapter 8 as well as a known experimental structure useful for comparison. The average intermolecular spacing values are 0.57 nm, 0.72 nm, and 0.76 nm considering 1, 2, and 3 neighbouring molecules, respectively. The individual distances between corannulene monomers within the crystal structure calculated in this analysis are in excellent agreement with those reported in Sanyal et al. [285]. Figure C.2 shows the crystal structure and alignment angles examined in Chapter 8.